ADVANCED CALCULUS

THE MACMILLAN COMPANY
NEW YORK · BOSTON · CHICAGO
DALLAS · ATLANTA · SAN FRANCISCO

MACMILLAN AND CO., LIMITED
LONDON · BOMBAY · CALCUTTA
MADRAS · MELBOURNE

THE MACMILLAN COMPANY
OF CANADA, LIMITED
TORONTO

ADVANCED CALCULUS

BY

WILLIAM F. OSGOOD, PH.D., LL.D.

PERKINS PROFESSOR OF MATHEMATICS
IN HARVARD UNIVERSITY

New York

THE MACMILLAN COMPANY

Set up and electrotyped. Published November, 19
Sixteenth Printing April, 1947.

PRINTED IN THE UNITED STATES OF AMERICA

PREFACE

Any course in Advanced Calculus must deal with partial differentiation and multiple integrals, with systematic integration, with improper integrals, and to some extent with complex quantities. Before modern methods in analysis — "ϵ-methods," as they are sometimes called — had been developed, the race became aware of a wide range of applications in physics (including geometry, the noblest branch of physics) which, although not yet technical, nevertheless make exacting demands on precise formulation and thus bring out both the physical hypotheses and the analytic means of working with them. The deduction of the partial differential equation which governs the flow of heat or electricity in conductors, the establishment of the equation of continuity in hydrodynamics and elasticity, and the setting up of the equations which describe the motion of the vibrating string or membrane, are cases in point. In these days when modern physics is primarily interested in the motion of discrete particles, it is particularly timely to emphasize continuous distributions of physical substances throughout regions of space, and continuous transformations of space.

The demands which geometry makes on partial differentiation are relatively slight. In thermodynamics a thoroughgoing appreciation of what the independent variables are (in order that, when the letters expressing the variables of two classes overlap, the meaning of the partial derivatives may be clear) and the ability to think in terms of line integrals, are indispensable.

Oscillatory motion is a basal conception in physics. Simple harmonic motion; next, the simplest case of damping; and finally the case of an impressed periodic force — these physical pictures are important alike for the student of physics and the student of pure mathematics, for they help to give him perspective as he proceeds with the study of the chapter in differential equations which relates to the solution of boundary value problems by means of developments into series. Moreover, Fourier's series and similar developments are illumined by the applications they here find.

In the early days of modern mathematics no sharp distinction was made between the Differential and Integral Calculus, and the Calculus of Variations — it was all, the Infinitesimal Calculus. With the knowledge of the great delicacy, both in concepts and methods, of which the Calculus of Variations is capable, came to mathematicians an awe of this subject, which resulted in a certain aloofness; the subject became a topic in the theory of functions of real variables. For the physicist, however, Hamilton's Principle is indispensable, and he has been obliged to get together some account of the rudiments of the Calculus of Variations as best he may. Sufficient conditions for a maximum or a minimum of an integral do not interest him. He needs to know when a certain integral is *stationary*, and this condition depends on the definition of a variation, δx, δU, etc. It is, therefore, essential that this definition be treated with care from the start, for it becomes increasingly complex as one proceeds. The Principle is applied to a variety of important problems in elastic vibrations.

There is a chapter on the systematic treatment of differential equations. But what is far more important is the unsystematic treatment of differential equations, which permeates these two volumes on the Calculus, beginning with the chapter on Mechanics in the *Introduction to the Calculus*. I have, moreover, taken occasion in the present chapter to point out the inner meaning of a differential equation through the geometric picture of a field of infinitesimal vectors or an assemblage of surface elements, and have thus led up to the idea of the integrals as families of curves, or of surfaces generated by characteristic strips.

As regards method, it sometimes happens that the naïve use of infinitesimals, even when it cannot be directly justified, has suggestive heuristic value; consider, for example, the transformation of multiple integrals and the flux across a surface; Chapter XII. In such cases, I have taken pains to conserve all that is helpful in these primitive conceptions, and have then supplemented them by proofs which meet our present standards of rigor. In this connection may also be cited (although it is not a question here of infinitesimals) the note on density and specific pressure or specific force; Chap. III, § 14.

A new form of the definition of a definite integral, simple or multiple, makes possible a simple and rigorous proof of the Fundamental Theorem of the Integral Calculus; Chap. XII, §§ 1–3.

There is a chapter on Vector Analysis, with applications to the proof of Stoke's Theorem and the deduction of the Frenet formulas.

Lagrange's Multipliers appear in maxima and minima of functions of several variables. Fourier's series and the allied developments into series of Bessel's functions and zonal harmonics are treated from the point of view of making the integral of the square of the error a minimum.

In the foregoing I have been describing those aims of the book which are not common in the text-books of the present day. To attain these ends, a purely mathematical treatment, availing itself of that which is best in the mathematics of today, but at the same time adapted to the powers (and the weaknesses) of the Junior or Senior in our colleges and schools of technology, must go before; and, indeed, not only the early parts of the various chapters, but by far the greater part of the space throughout the whole book is devoted to matters of an elementary nature. The book begins with the most rudimentary properties of polynomials and fractions, in preparation for integration, and the last chapter might well have been entitled: "The Story of $\sqrt{-1}$." It may seem exorbitant to spend ten pages on the study of integrals involving $\sqrt{a + bx + cx^2}$ and yet, a thoroughgoing understanding of all that is here involved covers substantially the whole field of systematic integration. But why should a physicist worry about the sign of a factor removed from under a radical sign? Merely because an error here gives him a wrong result in a problem on attractions.

The book is so written as to afford the greatest latitude in the order in which the various topics may be taken up. Thus the student may begin with the chapter on Partial Differentiation, or Double Integrals, or Differential Equations. Even within a chapter there is often a choice; *cf.* for example the foot-notes on p.p. 44 and 106. Personally, I should not wish to begin the course with Chapter I. For, although the subject is largely formal, testing the student's training in high school algebra and teaching him how to evaluate somewhat intricate integrals, the treatment should also serve to give him insight into the methods of algebra, and it should encourage him to become acquainted, for example, with the early chapters of Bôcher's *Algebra*.

It is assumed that B. O. Peirce's *A Short Table of Integrals*, Ginn & Co., Boston, is in the hands of the student. The references to *Analytic Geometry* are to Osgood and Graustein's *Plane and Solid Analytic Geometry*, Macmillan, 1921.

An honest attempt has been made to meet the undergraduate on his own ground. The appeal is not merely to the specialist in mathematics or physics; it is to all who would possess themselves of the Calculus as a method for understanding, in the broadest sense of the term, the quantitative relations which follow from the laws of nature.

CAMBRIDGE, MASSACHUSETTS,
September, 1925.

CONTENTS

CHAPTER I

GENERAL METHODS OF INTEGRATION

CHAPTER II

REDUCTION FORMULAS

CHAPTER III

DOUBLE INTEGRALS

CHAPTER IV

TRIPLE INTEGRALS

CHAPTER V

PARTIAL DIFFERENTIATION

CHAPTER VI

APPLICATIONS TO THE GEOMETRY OF SPACE

CHAPTER VII

TAYLOR'S THEOREM. MAXIMA AND MINIMA. LAGRANGE'S MULTIPLIERS

CHAPTER VIII

ENVELOPES

CHAPTER IX

ELLIPTIC INTEGRALS

CHAPTER X

INDETERMINATE FORMS

CHAPTER XI

LINE INTEGRALS AND GREEN'S THEOREM. FLOW OF HEAT

CHAPTER XII

TRANSFORMATION OF MULTIPLE INTEGRALS. EQUATION OF CONTINUITY

CHAPTER XIII

VECTOR ANALYSIS

CHAPTER XIV

DIFFERENTIAL EQUATIONS

CHAPTER XVII

THE CALCULUS OF VARIATIONS AND HAMILTON'S PRINCIPLE

CHAPTER XVIII

THERMODYNAMICS. ENTROPY

CHAPTER XIX

DEFINITE INTEGRALS AND THE GAMMA FUNCTION

CALCULUS

CHAPTER I

GENERAL METHODS OF INTEGRATION

In the first treatment of Integration,* various devices were set forth, whereby many integrals could be evaluated, but no general methods of integration were discussed. The object of the present chapter is to show how certain large and important classes of functions can be systematically integrated. The leading theorem is this, *that every rational function can be integrated in terms of the elementary functions.* Its proof depends on certain properties of polynomials and fractions, and so we begin with the discussion of these properties.

1. **Polynomials.** By a *polynomial* is meant a sum of monomial terms,

$$c_1x^{n_1} + c_2x^{n_2} + \cdots + c_rx^{n_r},$$

where the exponents are positive integers or zero. Such a sum can be written in the form:

$$(1) \qquad G(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n,$$

where the coefficients, a_k, do not depend on x, and n is a positive integer, or 0. In particular, a polynomial may reduce to a single term, as x^2 or $-x$ or c or 2 or 0.

If a_n is not 0, the *degree* of $G(x)$ is defined as n. Thus the polynomials

$$a^3 - x^3, \qquad -x, \qquad 5$$

are respectively of degree 3, 1, and 0. The polynomial 0 has no degree, and it is the only polynomial which has no degree.**

It is clear that the sum, the difference, and the product of two

* Cf. the author's *Introduction to the Calculus*, 1922, Chap. IX.

** Although the treatment here given is complete, the student will find it useful to read the first chapter of Bôcher's *Algebra.*

values of x, where $m \leq n$, then the corresponding coefficients must be respectively equal:

$$a_0 = b_0, \qquad a_1 = b_1, \quad \cdots \quad a_n = b_n, \qquad\qquad n = m.$$

For, the difference, $G(x) - g(x)$, vanishes identically.

EXERCISES

Factor each of the following polynomials:

1. $x^4 - a^4$.
2. $x^3 + a^3$.
3. $x^4 + a^2x^2$.
4. $x^4 - 2a^2x^2 + a^4$.
5. $x^4 + 2a^2x^2 + a^4$.
6. $x^2 + 5x + 6$.
7. $15x^5 - 12x^4 - 3x^3$.
8. $x^2 + 6x + 7$. *Ans.* $(x + 3 + \sqrt{2})(x + 3 - \sqrt{2})$.
9. $2x^2 + x - 7$.
10. $6x^4 - x^3 - 6x^2$.
11. $x^6 - a^6$.
12. $x^8 - a^8$.
13. $2x^3 - 3$.

2. Fractions. A fraction is the quotient of two polynomials,* $g(x)/G(x)$. The numerator, $g(x)$, may be any polynomial whatever. The denominator, $G(x)$, may be any polynomial but 0.

If the degree of $g(x)$ is less than that of $G(x)$, the fraction is called a *proper fraction*. For example:

$$\frac{6x}{4 + x^2}, \qquad \frac{12}{2 - 3x}.$$

In all other cases, the fraction is called an *improper fraction*. For example:

$$\frac{2x^2 + 1}{x + 3}, \qquad \frac{x + 1}{x}.$$

The fractions include the polynomials as particular cases, since $G(x)$ may $= 1$.

It may happen that numerator and denominator have a common factor; *e.g.*

$$\frac{x^2 - a^2}{x^3 - a^3}.$$

* The term *fraction* in elementary algebra is also applied to expressions like

$$\left(\frac{1}{x} - \frac{1}{a}\right) \bigg/ \left(\frac{1}{x} + \frac{1}{a}\right),$$

which can be reduced to a fraction as defined above. It is preferable henceforth to denote such expressions as *rational functions*; cf. § 8. Sometimes expressions like $x/\sqrt{a - x}$, or $(\sin x)/x$ are called fractions. This use of the word will not occur in this book, since it would lead only to confusion.

When all such factors have been divided out, the resulting fraction is said to be *in its lowest terms.* Thus the above fraction, when reduced to its lowest terms, becomes:

$$\frac{x+a}{x^2+ax+a^2}.$$

When a fraction is given, the first thing to do is, if necessary, to reduce it to its lowest terms, and we shall tacitly assume that this has been done.

An improper fraction (which is not a polynomial) can be reduced, by the process of division, to the sum of a polynomial and a proper fraction. For example,

$$\frac{2x^4-x^3-2x^2-3x-3}{x^2+x+1}=2x^2-3x-1+\frac{x-2}{x^2+x+1}.$$

Since our ultimate object is that of integrating a given fraction, *i.e.* of evaluating

$$\int\frac{g(x)}{G(x)}dx,$$

and since the integration of polynomials presents no difficulty, we shall be interested in the further study only of proper fractions.

EXERCISES

Reduce each of the following fractions to the sum of a polynomial and a proper fraction.

1. $\dfrac{x^2+x+1}{x}$. 2. $\dfrac{x^2+x+1}{x^2}$. 3. $\dfrac{x^2+x+1}{x+1}$.

4. $\dfrac{x^4+1}{(x-1)^2}$. 5. $\dfrac{x^4+1}{x^3+x^2}$. 6. $\dfrac{x^2-5x+6}{5x+2}$.

7. $\dfrac{x^4-8x^3+24x^2-32x+16}{x^2+x+1}$. 8. $\dfrac{x^5-7x+1}{(x-1)^3}$.

3. Partial Fractions. It is possible to express any proper fraction as the sum of fractions of the following types:

(*i*) $\dfrac{A}{x-a}$, $\dfrac{A}{(x-a)^m}$, m, a positive integer;

(*ii*) $\dfrac{Ax+B}{x^2+px+q}$, $\dfrac{Ax+B}{(x^2+px+q)^m}$, $p^2-4q<0$.

20. $\int \frac{(6x+13)\,dx}{6x^2-7x-3}$. 21. $\int \frac{(2-3x)\,dx}{3+2x-x^2}$. 22. $\int \frac{(x^2+1)\,dx}{3x^2+x-1}$.

4. Continuation. Multiple Linear Factors. If $G(x)$ contains the factor $x-a$ m times:

$$G(x) = (x-a)^m \phi(x), \qquad 1 \leqq m,$$

where $\phi(x)$ is not divisible by $x-a$, then the given fraction can be written in the form:

$$\frac{g(x)}{G(x)} = \frac{A_1}{(x-a)^m} + \frac{A_2}{(x-a)^{m-1}} + \cdots + \frac{A_m}{x-a} + \frac{f(x)}{\phi(x)}, \tag{1}$$

where $f(x)/\phi(x)$ is a proper fraction, or zero, and $A_1 \neq 0$. For the proof, cf. § 5.

Example 1.

$$\frac{-3x^2-2x+3}{x^3+x^2} = \frac{A}{x^2} + \frac{B}{x} + \frac{C}{x+1}; \tag{2}$$

for, since $\phi(x) = x+1$ is here of the first degree, $f(x)$ must reduce to a constant.

We can now proceed as in the last paragraph, clearing equation (2) of fractions:

$$-3x^2-2x+3 = A(x+1) + B(x^2+x) + Cx^2,$$

or

$$3-2x-3x^2 = A + (A+B)x + (B+C)x^2.$$

This equation will hold for all values of x if A, B, and C can be so determined that

$$A = 3, \qquad A+B = -2, \qquad B+C = -3.$$

Solving these equations, we find

$$A = 3, \qquad B = -5, \qquad C = 2.$$

Hence, on substituting these values and retracing our steps, we find:

$$\frac{-3x^2-2x+3}{x^3+x^2} = \frac{3}{x^2} - \frac{5}{x} + \frac{2}{x+1}.$$

The truth of this equation can be at once verified.

Example 2.

$$\frac{x}{x^3-1} = \frac{A}{x-1} + \frac{Bx+C}{x^2+x+1},$$

for here $\phi(x) = x^2+x+1$, and so $f(x)$ can be at most of the first degree. Proceeding as before, we have:

$$x = A(x^2 + x + 1) + (Bx + C)(x - 1),$$

$$x = A - C + (A - B + C)x + (A + B)x^2;$$

$$A - C = 0, \qquad A - B + C = 1, \qquad A + B = 0;$$

$$A = \tfrac{1}{3}, \qquad B = -\tfrac{1}{3}, \qquad C = \tfrac{1}{3};$$

$$\frac{x}{x^3 - 1} = \frac{1}{3(x-1)} + \frac{-x+1}{3(x^2+x+1)}.$$

Evaluation of the Integral $\displaystyle\int \frac{(Ax + B)\,dx}{x^2 + px + q}$, $p^2 - 4q < 0$.

When $p = 0$, q is positive and can be set $= a^2$, and the evaluation is immediate, for

(3) $$\int \frac{x\,dx}{x^2 + a^2} = \frac{1}{2}\log(x^2 + a^2),$$

(4) $$\int \frac{dx}{x^2 + a^2} = \frac{1}{a}\tan^{-1}\frac{x}{a}.$$

If $p \neq 0$, a linear transformation serves to reduce the given integral to the forms (3) and (4). We can write

$$x^2 + px + q = x^2 + px + \left(\frac{p}{2}\right)^2 + q - \frac{p^2}{4} = \left(x + \frac{p}{2}\right)^2 + a^2,$$

where the positive number, $q - p^2/4$, has been set equal to a^2. Next, set

$$t = x + \frac{p}{2};$$

$$\int \frac{dx}{x^2 + px + q} = \int \frac{dt}{t^2 + a^2} = \frac{1}{a}\tan^{-1}\frac{t}{a} = \frac{2}{\sqrt{4q - p^2}}\tan^{-1}\frac{2x + p}{\sqrt{4q - p^2}}.$$

Similarly:

$$\int \frac{x\,dx}{x^2 + px + q} = \int \frac{(t - \frac{1}{2}p)\,dt}{t^2 + a^2}.$$

This last integral comes back to (3) and (4), and thus the integration is accomplished.

Example. To evaluate $\displaystyle\int \frac{x\,dx}{x^2 + 4x + 9}$.

Here, $t = x + 2$, $a^2 = 5$,

$$\int \frac{x\,dx}{x^2 + 4x + 9} = \int \frac{(t - 2)\,dt}{t^2 + a^2}$$

$$= \tfrac{1}{2}\log(x^2 + 4x + 9) - \frac{2}{\sqrt{5}}\tan^{-1}\frac{x + 2}{\sqrt{5}} + C.$$

EXERCISES

Express the following fractions in terms of partial fractions.

1. $\dfrac{3x^2-2x-3}{x^3-x^2}$. 2. $\dfrac{6}{x^3-6x^2+9x}$. 3. $\dfrac{1}{(x^2-1)^2}$.

4. $\dfrac{a^2}{(x^2-a^2)^2}$. 5. $\dfrac{x^2-2x+1}{x^3}$. 6. $\dfrac{x^2-2}{(x+1)^3}$.

7. $\dfrac{x^2+x+1}{(x-2)^4}$. 8. $\dfrac{x}{x^3+1}$. 9. $\dfrac{2x^2-x-4}{x^3-1}$.

10. $\dfrac{4a^3}{x^4-a^4}$. *Ans.* $\dfrac{1}{x-a}-\dfrac{1}{x+a}-\dfrac{2a}{x^2+a^2}$.

11. $\dfrac{x}{x^4-a^4}$. 12. $\dfrac{x+4}{2x^4+x}$. 13. $\dfrac{x^6}{x^4-a^4}$.

14. $\dfrac{-2x^3+2x^2-2x+6}{(x-1)^2(x^2-x+2)}$. *Ans.* $\dfrac{2}{(x-1)^2}-\dfrac{3}{x-1}+\dfrac{x-4}{x^2-x+2}$.

15. $\dfrac{2x^3+2x^2+2x+6}{(x+1)^2(x^2+x+2)}$. 16. $\dfrac{x+1}{(x+2)^2(x^2+1)}$.

Evaluate the following integrals, using the *method* of the text, *not* the final formulas.

17. $\displaystyle\int\frac{dx}{x^2+2x+2}$. 18. $\displaystyle\int\frac{(5x-6)\,dx}{3x^2-2x+3}$. 19. $\displaystyle\int\frac{x^2\,dx}{x^2+2x+2}$.

20. Integrate the fraction of Question 14.

5. Lemma I. The proof of the theorem about partial fractions rests on two lemmas.

LEMMA I. *If $G(x)$ contains the factor $x-a$ precisely m times:*

$$G(x)=(x-a)^m\phi(x),\qquad 1\leq m,\qquad \phi(a)\neq 0,$$

the fraction can be written in the form:

$$\frac{g(x)}{G(x)}=\frac{A}{(x-a)^m}+\frac{f(x)}{(x-a)^{m-1}\phi(x)},\qquad A\neq 0, \tag{1}$$

where the last term is a proper fraction, or zero.

Form the difference,

$$\frac{g(x)}{G(x)}-\frac{A}{(x-a)^m},$$

where A is any constant whatever, and reduce to a common denominator:

$$\frac{g(x)}{G(x)} - \frac{A}{(x-a)^m} = \frac{g(x) - A\phi(x)}{(x-a)^m \phi(x)}.$$

If we can determine A so that the numerator of the last fraction is divisible by $x - a$, then we can cancel $x - a$ at least once from numerator and denominator, and on transposing the term $A/(x-a)^m$, the theorem is proved.

The condition that $x - a$ divide the numerator is that $x = a$ be a root, or

$$g(a) - A\phi(a) = 0.$$

This equation can always be solved for A:

$$(2) \qquad A = \frac{g(a)}{\phi(a)},$$

since by hypothesis $\phi(a) \neq 0$. This completes the proof.

The last fraction in (1) is not necessarily in its lowest terms, for it may happen that a higher power of $x - a$ can be cancelled from numerator and denominator. But no factor of $\phi(x)$ can divide $f(x)$. For, multiply (1) by $G(x)$:

$$g(x) = A\phi(x) + (x-a)f(x).$$

A factor common to $\phi(x)$ and $f(x)$ would thus divide $g(x)$. Hence $g(x)$ and $\phi(x)$ would have a common factor, and the original fraction, $g(x)/G(x)$, would not be in its lowest terms.

We observe that equation (2) gives an explicit determination of A, and thus avoids the computation of the earlier method, § 3, and, in a measure, the computation in § 4.

Example 1. Consider the fraction

$$\frac{g(x)}{G(x)} = \frac{x^2 + x + 1}{(x-1)(x-2)(x-3)}.$$

By the lemma, we can write it in the form

$$\frac{g(x)}{G(x)} = \frac{A}{x-1} + \frac{f(x)}{\phi(x)},$$

where $\quad a = 1, \qquad \phi(x) = (x-2)(x-3), \qquad g(x) = x^2 + x + 1.$

Hence
$$A = \frac{g(1)}{\phi(1)} = \frac{3}{(-1)(-2)} = \frac{3}{2},$$

and
$$\frac{x^2 + x + 1}{(x-1)(x-2)(x-3)} = \frac{3}{2(x-1)} + \frac{f(x)}{(x-2)(x-3)}.$$

We could compute $f(x)$ by transposing the first term on the right and reducing, and then applying the lemma to the new fraction. But the explicit determination of $f(x)$ is unnecessary. We see from the lemma that

$$\frac{f(x)}{(x-2)(x-3)}=\frac{B}{x-2}+\frac{C}{x-3},$$

and hence the original fraction can be written:

$$\frac{g(x)}{G(x)}=\frac{3}{2(x-1)}+\frac{B}{x-2}+\frac{C}{x-3}.$$

Now, we might equally well have begun with the root 2 of $G(x)$. Denote it by b; denote the corresponding function $\phi(x)$ by $\phi_1(x)$:

$$\phi_1(x)=(x-1)(x-3);$$

and denote the new A by B:

$$B=\frac{g(b)}{\phi_1(b)}=\frac{g(2)}{\phi_1(2)}=\frac{7}{-1}.$$

Thus B is determined: $B=-7$. C can be found in a precisely similar way; its value turns out to be $\frac{13}{2}$. Hence, finally,

$$\frac{x^2+x+1}{(x-1)(x-2)(x-3)}=\frac{3}{2(x-1)}-\frac{7}{x-2}+\frac{13}{2(x-3)}.$$

Uniqueness. There is really a question of uniqueness here involved. How do we know that the B determined in the one way and the B determined in the other are the same number?

Suppose they were not. Then we should have:

$$\frac{g(x)}{G(x)}=\frac{B}{x-2}+R(x) \qquad \text{and} \qquad \frac{g(x)}{G(x)}=\frac{B'}{x-2}+S(x),$$

where $R(x)$ and $S(x)$ are fractions, each continuous at $x=2$. Subtract the one equation from the other:

$$0=\frac{B-B'}{x-2}+R(x)-S(x).$$

And now let x approach the limit 2. $R(x)$ and $S(x)$ both approach limits. Hence $B-B'=0$; for otherwise the first term on the right would become infinite. Thus $B'=B$, $S(x)\equiv R(x)$, and the uniqueness is established.

In the same way it can be shown generally that the A of formula (1) is uniquely determined. This fact can be read off at once from the equation

$$g(a)-A\phi(a)=0,$$

since the latter represents a *necessary*, as well as a *sufficient*, condition that $g(x) - A\phi(x)$ be divisible by $x - a$.

Example 2. Let

$$\frac{g(x)}{G(x)} = \frac{3x-1}{x^4+x^2}.$$

Here, $\quad a = 0, \quad \phi(x) = x^2 + 1, \quad g(x) = 3x - 1,$

and

$$A = \frac{g(0)}{\phi(0)} = -1.$$

Thus

$$\frac{3x-1}{x^4+x^2} = -\frac{1}{x^2} + \frac{f(x)}{x^3+x}.$$

The next step is most conveniently taken by actually computing $f(x)$. This can be done by transposing, reducing, and multiplying through by $x^3 + x$. Thus we find: $f(x) = x + 3$. We now deal with the new fraction,

$$\frac{x+3}{x^3+x},$$

setting $\quad a = 0, \quad \phi(x) = x^2 + 1, \quad g(x) = x + 3,$

and thus see that $A = g(0)/\phi(0) = 3$; hence

$$\frac{x+3}{x^3+x} = \frac{3}{x} + \frac{Cx+D}{x^2+1}.$$

The coefficients, C and D, could be obtained by clearing of fractions and comparing coefficients. A shorter method, however, is the following. First, multiply through by x:

$$\frac{x+3}{x^2+1} = 3 + \frac{Cx^2+Dx}{x^2+1};$$

and now allow x to become infinite. Thus we see that $C = -3$, and

$$\frac{x+3}{x^3+x} = \frac{3}{x} + \frac{-3x+D}{x^2+1}.$$

Finally, to determine D, give to x any special value for which no denominator vanishes. A simple value is $x = 1$; thus

$$2 = 3 + \frac{-3+D}{2}, \qquad D = 1,$$

and the final result is:

$$\frac{3x-1}{x^4+x^2} = -\frac{1}{x^2} + \frac{3}{x} - \frac{3x-1}{x^2+1}.$$

The truth of this equation can be verified by reducing the right-hand side to a common denominator.

We now have at our disposal (*i*) the method of undetermined coefficients, set forth in §§ 3, 4; (*ii*) the explicit formula (2); (*iii*) the method of giving to x a number of convenient special values, or of allowing x to become infinite, after multiplying the equation through by a suitable power of x. The best results are obtained by a skillful combination of all these methods.

EXERCISES

Work a number of the Exercises of § 4 by means of the present labor-saving devices.

6. Lemma II. We proceed now to the second lemma.

Lemma II. *If $G(x)$ contains the factor x^2+px+q precisely m times:*

$$G(x)=(x^2+px+q)^m\phi(x), \quad 1 \leqq m, \quad p^2-4q<0, \left\{\begin{array}{l}\phi(x) \text{ not divisible by} \\ x^2+px+q.\end{array}\right.$$

the fraction can be written in the form:

$$(1) \qquad \frac{g(x)}{G(x)}=\frac{Ax+B}{(x^2+px+q)^m}+\frac{f(x)}{(x^2+px+q)^{m-1}\phi(x)}, \qquad \left\{\begin{array}{l}A \text{ and } B \\ \text{not both } 0.\end{array}\right.$$

where the last term is a proper fraction, or zero.

Form the difference, and reduce to a common denominator:

$$(2) \qquad \frac{g(x)}{G(x)}-\frac{Ax+B}{(x^2+px+q)^m}=\frac{g(x)-(Ax+B)\phi(x)}{(x^2+px+q)^m\phi(x)}.$$

We wish to show that A and B can be so determined that the numerator is divisible by x^2+px+q. Divide $g(x)$ and $\phi(x)$ by x^2+px+q:

$$g(x)=Q_1(x)(x^2+px+q)+Lx+M,$$
$$\phi(x)=Q_2(x)(x^2+px+q)+\lambda x+\mu.$$

Here, one or both of the quotients, $Q_1(x)$ and $Q_2(x)$, may be 0, — that makes no difference. But L and M cannot both vanish, for then $g(x)$ would be divisible by x^2+px+q, and so the original fraction, $g(x)/G(x)$, would not be in its lowest terms. And similarly, λ and μ cannot both be 0, for then $\phi(x)$ would be divisible by x^2+px+q. — contrary to hypothesis.

If, now, the numerator on the right of (2) is to be divisible by x^2+px+q, it is clearly necessary and sufficient that

$$Lx+M-(Ax+B)(\lambda x+\mu)$$

be divisible by x^2+px+q. Multiply out:

$$Lx+M-(Ax+B)(\lambda x+\mu)=-\lambda Ax^2-(\mu A+\lambda B-L)x-\mu B+M;$$

and now divide this last polynomial by x^2+px+q:

$$\begin{array}{r|l} & -\lambda A \\ x^2+px+q & -\lambda Ax^2-(\mu A+\lambda B-L)x-\mu B+M \\ & -\lambda Ax^2 \qquad\qquad -p\lambda Ax-q\lambda A \\ \hline & \{(-\mu+p\lambda)A-\lambda B+L\}x+q\lambda A-\mu B+M \end{array}$$

The last line gives the remainder, and this must vanish identically. Hence we must have:

$$(3)\qquad \begin{cases} (-\mu+p\lambda)A-\lambda B+L=0, \\ q\lambda A-\mu B+M=0. \end{cases}$$

These are two linear equations for determining the unknowns, A and B. They are non-homogeneous, since L and M are not both 0. Their determinant is:

$$\begin{vmatrix} -\mu+p\lambda & -\lambda \\ q\lambda & -\mu \end{vmatrix}=\mu^2-p\lambda\mu+q\lambda^2.$$

Its value is not 0. For, we can write it in the form:

$$\mu^2-p\lambda\mu+q\lambda^2=\lambda^2\left\{\left(-\frac{\mu}{\lambda}\right)^2+p\left(-\frac{\mu}{\lambda}\right)+q\right\},$$

provided $\lambda\neq 0$. Now, the brace cannot vanish; for then the quadratic polynomial x^2+px+q would have a real root, $x=-\mu/\lambda$. If, on the other hand, $\lambda=0$, the above determinant reduces to μ^2. Since λ and μ are not both 0, the determinant does not vanish in this case, either.

Hence equations (3) admit, in all cases, one and only one solution, and A and B are uniquely determined. They are not both 0, for then L and M would both be 0. This completes the proof.

It may happen that the last fraction in (1) is not in its lowest terms, the numerator being divisible by a power of x^2+px+q. But no factor of $\phi(x)$ can divide the numerator.

Example 1. Let

$$\frac{g(x)}{G(x)}=\frac{x^2}{(x^2+x+1)(x^2+1)}.$$

Then we have:

$$\frac{x^2}{(x^2+x+1)(x^2+1)}=\frac{Ax+B}{x^2+x+1}+\frac{Cx+D}{x^2+1}.$$

First, multiply through by x:

$$\frac{x^3}{(x^2+x+1)(x^2+1)}=\frac{Ax^2+Bx}{x^2+x+1}+\frac{Cx^2+Dx}{x^2+1}.$$

Now allow x to become infinite, and we find:

i) $$A+C=0.$$

Next, give to x the simplest possible value, namely 0:

ii) $$B+D=0.$$

The next simplest values for x are 1 and -1:

iii) $\frac{1}{6}=\frac{A+B}{3}+\frac{C+D}{2}$, or $2A+2B+3C+3D=1$.

iv) $\frac{1}{2}=-A+B+\frac{-C+D}{2}$, or $-2A+2B-C+D=1.$

From (*i*) and (*ii*):

$$C=-A, \qquad D=-B.$$

Substituting these values for C and D in (*iii*) and (*iv*), we get:

$$-A-B=1,$$
$$-A+B=1.$$

Hence $A=-1$, $B=0$. Moreover, $C=1$, $D=0$. Thus

$$\frac{x^2}{(x^2+x+1)(x^2+1)}=\frac{-x}{x^2+x+1}+\frac{x}{x^2+1}.$$

Example 2. Let

$$\frac{g(x)}{G(x)}=\frac{x^3}{(x^2+x+1)^2}.$$

Here, $\phi(x)$ is of the 0-th degree, and it is simpler not to use the lemma, but to proceed directly by division:

$$\begin{array}{r|l} & x-1 \\ x^2+x+1 & x^3 \\ & \underline{x^3+x^2+x} \\ & -x^2-x \\ & \underline{-x^2-x-1} \\ & 1 \end{array}$$

Hence

$$\frac{x^3}{x^2+x+1}=x-1+\frac{1}{x^2+x+1},$$

and

$$\frac{x^3}{(x^2+x+1)^2}=\frac{1}{(x^2+x+1)^2}+\frac{x-1}{x^2+x+1}.$$

This method is also simpler in such cases as

$$\frac{g(x)}{G(x)}=\frac{4x^2-x+5}{(2x-3)^3}.$$

EXERCISES

Express the following fractions in terms of partial fractions.

1. $\dfrac{x^3-x^2}{(2x^2-x+2)(x^2-2x+3)}$. 2. $\dfrac{1}{x^5+2x^3+x}$. 3. $\dfrac{x}{x^4+a^4}$.

4. $\dfrac{1}{(x^3+x^2+x)^2}$. 5. $\dfrac{x^3-x^2}{(2x^2-x+5)^2}$. 6. $\dfrac{x^5-a^5}{(x^2+a^2)^3}$.

7. Proof of the Theorem on Partial Fractions. Let $g(x)/G(x)$ be a proper fraction in its lowest terms. If

$$G(x)=(x-a)^m\phi(x), \qquad 1\leqq m,$$

where $\phi(x)$ is not divisible by $x-a$, then the fraction can, by the repeated application of Lemma I, be written in the form:

$$\frac{g(x)}{G(x)}=\frac{A_1}{(x-a)^m}+\frac{A_2}{(x-a)^{m-1}}+\cdots+\frac{A_m}{x-a}+\frac{\omega(x)}{\phi(x)},$$

where the last term on the right is either a proper fraction in its lowest terms, or zero. A_1 cannot vanish; but any or all the later A's may be zero.

Similarly, if

$$G(x)=(x^2+px+q)^m\phi(x), \qquad 1\leqq m, \qquad p^2-4q<0,$$

where $\phi(x)$ is not divisible by x^2+px+q, the fraction can, by the repeated application of Lemma II, be written in the form:

$$\frac{g(x)}{G(x)}=\frac{A_1x+B_1}{(x^2+px+q)^m}+\frac{A_2x+B_2}{(x^2+px+q)^{m-1}}+\cdots+\frac{\omega(x)}{\phi(x)},$$

where the last term is a proper fraction in its lowest terms, or zero. A_1 and B_1 cannot both be zero; the later A's and B's are subject to no restriction.

Since in each case the degree of $\phi(x)$ is less than that of $G(x)$, it is clear that sufficient repetitions of the above processes will finally reduce the original fraction to a sum of partial fractions, and the theorem is proved. The representation is unique.

8. Integration of Rational Functions. By a *rational function* of x is meant a fraction, or a so-called "complex fraction," like

$$\frac{1}{x} - 2x + ax^2 - \frac{\frac{x}{a} - \frac{a}{x}}{\frac{x}{a} + \frac{a}{x}}.$$

A rational function, $R(x)$, can, therefore, always be reduced to an ordinary fraction:*

$$R(x) = \frac{g(x)}{G(x)},$$

and hence it can be represented either as a polynomial or as a proper fraction or as the sum of a polynomial and a proper fraction.

We can now show that the integral of a rational function:

$$\int R(x)\,dx,$$

can always be evaluated in terms of the elementary functions. For, the polynomial part presents no difficulty, and the fractional part can be expressed in terms of partial fractions. The latter can be integrated as follows.

The integrals of the types

$$\int \frac{dx}{(x-a)^m}, \quad 1 \leqq m; \qquad \int \frac{(Ax+B)\,dx}{x^2+px+q}, \quad p^2 - 4q < 0,$$

are familiar to us. There remain only the integrals

$$\int \frac{(Ax+B)\,dx}{(x^2+px+q)^m}, \quad 1 < m, \quad p^2 - 4q < 0.$$

* Similarly, a rational function of two or more variables is any expression that can be put together out of these variables by means of the *four species*, — addition, subtraction, multiplication, and division, — *i.e.* it is a polynomial, or a "simple" or "complex" fraction. It can always be reduced to a polynomial or an ordinary fraction. Thus, for two variables,

$$R(x, y) = \frac{g(x, y)}{G(x, y)},$$

where $g(x, y)$ and $G(x, y)$ are polynomials.

To deal with these latter we make first the substitution of § 4:

$$t = x + \frac{p}{2}, \qquad a^2 = q - \frac{p^2}{4}.$$

Thus the above integral is reduced to a linear combination of the two integrals:

$$\int \frac{t\,dt}{(t^2 + a^2)^m}, \qquad \int \frac{dt}{(t^2 + a^2)^m}.$$

The first of these integrals can be found by the substitution

$$z = t^2 + a^2.$$

The second is obtained by a reduction formula:*

$$\int \frac{dt}{(t^2 + a^2)^m} = \frac{t}{(2m-2)\,a^2 (t^2 + a^2)^{m-1}} + \frac{2m-3}{(2m-2)\,a^2} \int \frac{dt}{(t^2 + a^2)^{m-1}}.$$

On replacing t by its value in terms of x this formula reduces substantially to Formula 71 of Peirce's *Tables*.

Thus the evaluation is complete and the theorem is proved.

9. The Integral $\int R(\sin x, \cos x)\,dx$.

By means of the theorem just established it can be shown that any rational function of $\sin x$ and $\cos x$ can be integrated. Make the substitution

(1) $$t = \tan \frac{x}{2}, \qquad x = 2 \tan^{-1} t, \qquad -\pi < x < \pi.$$

Then $$dx = \frac{2\,dt}{1 + t^2},$$

$$\sin x = 2 \sin \frac{x}{2} \cos \frac{x}{2} = 2 \tan \frac{x}{2} \cos^2 \frac{x}{2} = \frac{2 \tan \frac{1}{2} x}{\sec^2 \frac{1}{2} x},$$

or

(2) $$\sin x = \frac{2t}{1 + t^2}, \qquad \text{and} \qquad \cos x = \frac{1 - t^2}{1 + t^2}.$$

On substituting these values for X and Y in $R(X, Y)$ the result is seen to be a rational function of t:

$$R(\sin x, \cos x) = r(t).$$

Hence $\int R(\sin x, \cos x)\,dx = \int r(t) \frac{2\,dt}{1 + t^2}$, *i.e.* the integral of a rational function of t.

* Reduction formulas play an important rôle in systematic integration, and it is well to treat them as an independent subject. At the present moment the student needs only § 2, in Chap. II, which he now should read.

Example. Consider the integral

$$\int \frac{dx}{a + b\cos x}, \qquad |a| \neq |b|. \tag{3}$$

On making the substitution (1) and reducing, we have:

$$\int \frac{2\,dt}{a + b + (a - b)t^2}.$$

There are two cases, according as $a + b$ and $a - b$ have the same sign or opposite signs.*

Case I. $0 < \dfrac{a+b}{a-b} = A^2, \qquad -\pi < x < \pi.$

$$\int \frac{2\,dt}{a + b + (a-b)t^2} = \frac{2}{a-b}\int \frac{dt}{t^2 + A^2} = \frac{2}{(a-b)A}\tan^{-1}\frac{t}{A},$$

$$\int \frac{dx}{a + b\cos x} = \frac{2}{(a-b)A}\tan^{-1}\left\{\frac{\tan\frac{1}{2}x}{A}\right\}. \tag{4}$$

If $a - b$ is positive, the formula can be written:

$$\int \frac{dx}{a + b\cos x} = \frac{2}{\sqrt{a^2 - b^2}}\tan^{-1}\left(\sqrt{\frac{a-b}{a+b}}\tan\frac{x}{2}\right) \qquad \begin{cases} a + b > 0, \\ a - b > 0. \end{cases} \tag{5}$$

But this last formula is false when $a + b$ and $a - b$ are both negative, for then

$$(a - b)\sqrt{\frac{a+b}{a-b}} = -\sqrt{a^2 - b^2}.$$

A form which is general, comprising all integrals which can occur under Case I, is the following:

$$\int \frac{dx}{a + b\cos x} = \frac{2}{\sqrt{a^2 - b^2}}\tan^{-1}\frac{\sqrt{a^2 - b^2}\tan\frac{1}{2}x}{a + b}. \tag{6}$$

Case II. $0 > \dfrac{a+b}{a-b} = -A^2, \qquad -\pi < x < \pi.$

$$\int \frac{2\,dt}{a + b + (a-b)t^2} = \frac{2}{a-b}\int \frac{dt}{t^2 - A^2} = \frac{1}{(a-b)A}\log\frac{t - A}{t + A}.$$

$$\int \frac{dx}{a + b\cos x} = \frac{1}{(a-b)A}\log\frac{\tan\frac{1}{2}x - A}{\tan\frac{1}{2}x + A}. \tag{7}$$

If $a + b$ is positive and $a - b$, negative, the formula can be written·

* The cases in which a and b are numerically equal are dealt with directly.

$$(8)\quad \int\frac{dx}{a+b\cos x}=\frac{1}{\sqrt{b^2-a^2}}\log\frac{\sqrt{b+a}+\sqrt{b-a}\tan\frac{1}{2}x}{\sqrt{b+a}-\sqrt{b-a}\tan\frac{1}{2}x},\quad \begin{cases} b+a>0,\\ b-a>0.\end{cases}$$

But this formula is incomplete, failing to include, for example, the case $a=4$, $b=-5$. The following form is general:*

$$(9)\quad \int\frac{dx}{a+b\cos x}=\frac{1}{\sqrt{b^2-a^2}}\log\frac{a+b+\sqrt{b^2-a^2}\tan\frac{1}{2}x}{a+b-\sqrt{b^2-a^2}\tan\frac{1}{2}x}.$$

EXERCISES

Evaluate the following integrals, using each time the *method* of the text, not the final formula. Check by the formula.

1. $\int\frac{dx}{5+3\cos x}$.
2. $\int\frac{dx}{3+5\cos x}$.
3. $\int\frac{dx}{4\cos x-5}$.
4. $\int\frac{dx}{4-5\cos x}$.
5. $\int\frac{dx}{5+4\sin x}$.
6. $\int\frac{dx}{12\sin x-13}$.
7. $\int\frac{dx}{1-\cos x}$.
8. $\int\frac{dx}{1+\cos x}$.
9. $\int\frac{dx}{1-\sin x}$
10. $\int\frac{dx}{3-2\sin x+\cos x}$.
11. $\int\frac{dx}{1+\sin x+\cos x}$.
12. $\int\frac{dx}{1-\sin x+2\cos x}$.
13. $\int\frac{dx}{2-5\sin x+3\cos x}$.
14. $\int\frac{dx}{\cos x-\cos\alpha}$.
15. $\int\frac{dx}{\sin\alpha-\sin x}$.
16. $\int\frac{dx}{1+\cos\alpha\cos x}$.
17. Show that

$$\int\frac{dx}{a+b\sin x}=\begin{cases}\dfrac{2}{\sqrt{a^2-b^2}}\tan^{-1}\dfrac{b+a\tan\frac{1}{2}x}{\sqrt{a^2-b^2}}\\[2ex] \dfrac{1}{\sqrt{b^2-a^2}}\log\dfrac{b-\sqrt{b^2-a^2}+a\tan\frac{1}{2}x}{b+\sqrt{b^2-a^2}+a\tan\frac{1}{2}x}\end{cases}$$

according to which formula gives real results. Obtain formulas for the exceptional cases.

* Cf., however, § 3, *footnote*. In all our formulas of integration the inverse trigonometric functions are restricted to the principal values.

18. Show that

$$\int \frac{dx}{a + b\cos x + c\sin x} = \frac{2}{\sqrt{a^2 - b^2 - c^2}} \tan^{-1} \frac{c + (a - b)\tan \frac{1}{2}x}{\sqrt{a^2 - b^2 - c^2}}$$

or
$$\frac{1}{\sqrt{b^2 + c^2 - a^2}} \log \frac{\sqrt{b^2 + c^2 - a^2} - c - (a - b)\tan \frac{1}{2}x}{\sqrt{b^2 + c^2 - a^2} + c + (a - b)\tan \frac{1}{2}x},$$

according to which one of these formulas has a meaning. If neither has a meaning, obtain a formula that does have one.

19. Show that

$$\int \frac{dx}{a(1 + \cos x) + c\sin x} = \frac{1}{c}\log\left(a + c\tan\frac{x}{2}\right).$$

20. Evaluate by the present method

$$\int \sec x\, dx \qquad \text{and} \qquad \int \csc x\, dx.$$

21. By the aid of the identity

$$a\sin x + b\cos x = A\cos(x - \alpha),$$

where $\qquad A = \sqrt{a^2 + b^2}, \qquad \sin\alpha = a/A, \qquad \cos\alpha = b/A,$

obtain formulas for the integral of Question 18.

10. Integration by Ingenious Devices. Because every rational function of $\sin x$ and $\cos x$ can be integrated by the method of the last paragraph, it does not follow that that method is the simplest one in a given case. There may be an obvious substitution or reduction which leads at once to the result.

Example 1. $\qquad \displaystyle\int \frac{\sin x\, dx}{a + b\cos x}.$

Here, the numerator is substantially the differential of the denominator:

$$d(a + b\cos x) = -b\sin x\, dx,$$

and so the substitution

$$z = a + b\cos x$$

leads immediately to the desired evaluation.

Example 2. $\qquad \displaystyle\int \frac{\cos x\, dx}{a + b\cos x}.$

Divide the numerator by the denominator according to the methods of first-year algebra:

$$\frac{\cos x}{a+b\cos x}=\frac{1}{b}-\frac{a}{b(a+b\cos x)}.$$

Thus the integral is reduced to the integral studied in the text.

Example 3. $\int \cos^3 x\,dx.$

Here, $\cos^3 x\,dx = \cos^2 x\cos x\,dx = (1-\sin^2 x)\,d\sin x,$

and so the substitution $z=\sin x$ avails.

Example 4. $\int \frac{dx}{3\sin x - 4\cos x}.$

$$3\sin x - 4\cos x = 5(\tfrac{3}{5}\sin x - \tfrac{4}{5}\cos x).$$

Let $\cos\alpha = \tfrac{3}{5}, \qquad \sin\alpha = \tfrac{4}{5}.$

Thus α is completely determined, and

$$3\sin x - 4\cos x = 5\sin(x-\alpha).$$

The integral is hereby reduced to

$$\int \csc\phi\,d\phi.$$

It is not, however, merely to trigonometic cases that these remarks apply. Consider

Example 5. $\int \frac{x^2\,dx}{x^3+a^3}.$

Here, the method of partial fractions would be absurd, for

$$d(x^3+a^3)=3x^2\,dx,$$

and the substitution $z=x^3+a^3$ avails.

Example 6. $\int \frac{(2x^3-3x)\,dx}{x^6+x^2+1}.$

Substitute $z=x^2.$

The student should now turn back to the paragraph on Integration by Parts, *Introduction to the Calculus*, p. 243, and study it thoughtfully, working the examples afresh, in order to realize both the possibilities and the limitations of that method.

EXERCISES

1. $\int \frac{\cos^2 x\,dx}{2 - \cos x}$.

2. $\int \frac{\cos x\,dx}{7 + 13\sin x}$.

3. $\int \frac{\sin x \cos x\,dx}{9 - \cos x}$.

4. $\int \frac{\sin x \cos x\,dx}{4\sin x + 11}$.

5. $\int \frac{\sin x - \cos x}{\cos x - \cos \alpha}\,dx$.

6. $\int \frac{2\cos x - 3\sin x}{\sin \beta - \sin x}\,dx$.

7. $\int \frac{dx}{2\sin^2 x - 3\cos^2 x}$.

8. $\int \frac{dx}{A^2\cos^2 x + B^2\sin^2 x}$.

9. $\int \frac{2\,dx}{3\cos 2x - 2}$.

10. $\int \frac{dx}{\sin\frac{\alpha}{2} - \sin\frac{x}{2}}$.

11. $\int \sin^3 x\,dx$.

12. $\int \cos^5 x\,dx$

13. $\int \frac{dx}{\cos^3 x}$.

14. $\int \frac{dx}{\cos^2 x}$.

15. $\int \frac{dx}{\sin^3 x}$.

16. $\int \frac{dx}{\sin\frac{x}{2}}$.

17. $\int \frac{x^3\,dx}{1 + x^4}$.

18. $\int \frac{x\,dx}{5 - 3x^2 - x^4}$.

19. $\int \frac{x^5\,dx}{a^6 - x^6}$.

20. $\int \frac{dx}{1 - e^x}$.

21. $\int \frac{dx}{e^x + e^{-x}}$.

22. $\int \frac{e^x\,dx}{e^x + e^{-x}}$.

23. $\int e^{-ax}\cos bx\,dx$.

24. $\int e^{-ax}\sin bx\,dx$.

25. $\int x\cos x\,dx$

26. $\int xe^{-x^2}\,dx$.

27. $\int e^{-\cos x}\sin x\,dx$.

28. $\int \frac{\log x\,dx}{x^2}$.

29. $\int \log(a^2 - x^2)\,dx$.

30. $\int a\log\sqrt{a^2 + x^2}\,dx$.

31. $\int \frac{\log x\,dx}{(1 + x)^2}$.

32. $\int \frac{dx}{x\log x}$.

33. $\int \frac{\log x\,dx}{x}$.

34. $\int \csc^{-1} x\,dx$.

11. The Integral $\int R(x, \sqrt{a + bx + cx^2})\,dx$.

The integral of any rational function of x and the square root of a quadratic polynomial,

(1) $$y = \sqrt{a + bx + cx^2}, \qquad c \neq 0,^*$$

can be evaluated in terms of the elementary functions. For,

(2) $$a + bx + cx^2 = c\left\{\left(x + \frac{b}{2c}\right)^2 + \frac{4ac - b^2}{4c^2}\right\} = c(t^2 \pm A^2),$$

where $$t = x + \frac{b}{2c} \qquad \text{and} \qquad A^2 = \pm\frac{4ac - b^2}{4c^2},$$

the lower sign being chosen when $4ac - b^2$ is negative.

There are in all three cases to be distinguished.

Case I: $4ac - b^2 < 0, \quad c < 0.$ Here,

$$y = \sqrt{-c}\,\sqrt{A^2 - t^2},$$

and the substitution:

$$t = A\sin\theta \qquad (\text{ or } t = A\cos\theta\ )$$

reduces the integral to the form treated in § 9.

Example 1. $$\int\frac{dx}{\sqrt{a + bx + cx^2}}, \qquad c < 0.$$

Here, we must necessarily have $4ac - b^2 < 0$, since otherwise the radicand would take on no positive values.

$$t = x + \frac{b}{2c}, \qquad A = \frac{\sqrt{b^2 - 4ac}}{-2c} > 0, \qquad t = A\sin\theta;$$

$$dx = dt = A\cos\theta\,d\theta, \qquad \sqrt{a + bx + cx^2} = \sqrt{-c}\,A\cos\theta;$$

$$\int\frac{dx}{\sqrt{a + bx + cx^2}} = \int\frac{d\theta}{\sqrt{-c}} = \frac{1}{\sqrt{-c}}\sin^{-1}\frac{-2cx - b}{\sqrt{b^2 - 4ac}}.$$

The solution could, however, have been abbreviated by writing the integral in the form:

$$\frac{1}{\sqrt{-c}}\int\frac{d\left(x + \frac{b}{2c}\right)}{\sqrt{\frac{b^2 - 4ac}{4c^2} - \left(x + \frac{b}{2c}\right)^2}} = \frac{1}{\sqrt{-c}}\sin^{-1}\frac{x + \frac{b}{2c}}{\sqrt{\frac{b^2 - 4ac}{4c^2}}}.$$

* If $c = 0$ and $b \neq 0$, the same substitution, (1), is applicable, the rationalization of the integrand now being immediate:

$$y^2 = a + bx, \qquad x = \frac{y^2 - a}{b}, \qquad dx = \frac{2y\,dy}{b},$$

and the new variable of integration being y.

Let $$p = y - 2, \qquad q = x - 1.$$

Then the equation of the pencil will be:

(2) $$p - \lambda q = 0 \qquad \text{or} \qquad y - 2 = \lambda(x - 1).$$

A variable line of this pencil cuts the curve in the fixed point A and a second, variable point $P:(x, y)$. The coordinates of P are found in the usual way by eliminating y between (1) and (2):

(3) $$\sqrt{1 + x + 2x^2} = \lambda(x - 1) + 2,$$

$$1 + x + 2x^2 = \lambda^2(x - 1)^2 + 4\lambda(x - 1) + 4,$$

(4) $$-3 + x + 2x^2 = \lambda^2(x - 1)^2 + 4\lambda(x - 1).$$

We know that one root of equation (4) must be $x = 1$, since every line (2) goes through the point $A:(1, 2)$, and hence it should be possible to separate out a factor $x - 1$. It is at once obvious that this can be done, since

$$-3 + x + 2x^2 = (x - 1)(2x + 3).$$

On dividing (4) through by $x - 1$, we have:

(5) $$2x + 3 = \lambda^2(x - 1) + 4\lambda.$$

Hence

(6) $$x = \frac{-3 + 4\lambda - \lambda^2}{2 - \lambda^2}.$$

The value of y is found from (2) by substituting this value of x:

(7) $$y = \frac{4 - 5\lambda + 2\lambda^2}{2 - \lambda^2}.$$

Thus x and y have both been expressed as rational functions of a parameter λ; the conic has been *rationalized*. Conversely, λ can be expressed rationally in series of the coordinates of $P:(x, y)$, for from (2) we have

(8) $$\lambda = \frac{y - 2}{x - 1}.$$

The method is general, and can be applied to any conic, the vertex of the pencil being taken at an arbitrary point on the conic. In the case of a hyperbola, the pencil may be the lines parallel to an asymptote.*

* The method of rationalization does not in general apply to algebraic curves, $G(x, y) = 0$, of degree higher than the second. Those curves to which it does apply are called *unicursal*.

We can now evaluate the given integral.

$$(9) \qquad x-1=\frac{-5+4\lambda}{2-\lambda^2}, \qquad dx=\frac{2(4-5\lambda+2\lambda^2)\,d\lambda}{(2-\lambda^2)^2},$$

$$(10) \qquad \int\frac{dx}{(x-1)\sqrt{1+x+2x^2}}=\int\frac{2\,d\lambda}{-5+4\lambda}=\tfrac{1}{2}\log(4\lambda-5)$$

$$=\tfrac{1}{2}\log\frac{4\sqrt{1+x+2x^2}-5x-3}{x-1}.$$

The General Case:

$$\int\frac{dx}{(x-\rho)\sqrt{a+bx+cx^2}}, \qquad 0<a+b\rho+c\rho^2,$$

can be treated in the same manner. The curve

$$(1') \qquad y=\sqrt{a+bx+cx^2}$$

is cut by the line $x-\rho=0$ because of the condition of inequality imposed. It will be convenient to introduce the notation:

$$f(x)=a+bx+cx^2.$$

Then * $\qquad f(\rho)>0, \qquad f'(\rho)=b+2c\rho.$

Cut the curve (1′) by the pencil

$$(2') \qquad p-\lambda q=0, \qquad \text{or} \qquad y-\sqrt{f(\rho)}=\lambda(x-\rho).$$

The computation is parallel to that in the numerical case above considered. We find:

$$x=\frac{-b-c\rho+2\sqrt{f(\rho)}\lambda-\rho\lambda^2}{c-\lambda^2}, \qquad y=\frac{c\sqrt{f(\rho)}-f'(\rho)\lambda+\sqrt{f(\rho)}\,\lambda^2}{c-\lambda^2}.$$

$$\lambda=\frac{\sqrt{a+bx+cx^2}-\sqrt{f(\rho)}}{x-\rho}, \qquad dx=\frac{2[c\sqrt{f(\rho)}-f'(\rho)\lambda+\sqrt{f(\rho)}\,\lambda^2]d\lambda}{(c-\lambda^2)^2},$$

$$(10') \qquad \int\frac{dx}{(x-\rho)\sqrt{a+bx+cx^2}}$$

$$=\frac{1}{\sqrt{f(\rho)}}\log\frac{\sqrt{f(\rho)}\sqrt{a+bx+cx^2}-\tfrac{1}{2}f'(\rho)(x-\rho)-f(\rho)}{x-\rho},$$

where $f(\rho)=a+b\rho+c\rho^2>0.$

* The expression $b+2c\rho$ occurs later in the computation, and can be abbreviated as $f'(\rho)$. No property of the derivative is here involved. It appears to be merely an accident that the abbreviation is possible.

This formula can also be derived from the result of Ex. 2, § 11, by means of the substitution $x-\rho=1/t$. The method is not, however, so simple as would appear at first sight, since two cases must be distinguished when t^2 is taken out from under the radical, according as $t>0$ or $t<0$.

Example 2. Consider the same integral for the case that the conic is not cut by the line $x-\rho=0$. Here, c may be positive or negative; the conic, a hyperbola or an ellipse; but this conic must cut the axis of x in two points, $x=\alpha$ and $x=\beta$. Let $\alpha<\beta$.

To rationalize the conic, choose a pencil of lines with its vertex in the point $(\alpha, 0)$:

$$(11) \qquad p-\lambda q=0, \qquad \text{or} \qquad y=\lambda(x-\alpha).$$

Suppose that $c<0$. Then the conic is an ellipse, and x lies between α and β. We have:

$$a+bx+cx^2=(-c)(x-\alpha)(\beta-x),$$

where each factor on the right is positive. Thus

$$\sqrt{-c(x-\alpha)(\beta-x)}=\lambda(x-\alpha), \qquad -c(\beta-x)=\lambda^2(x-\alpha),$$

$$x=\frac{\alpha\lambda^2-c\beta}{\lambda^2-c}, \qquad y=\frac{-c(\beta-\alpha)\lambda}{\lambda^2-c};$$

$$\lambda=\sqrt{\frac{-c(\beta-x)}{x-\alpha}}; \qquad dx=\frac{2c(\beta-\alpha)\lambda\,d\lambda}{(\lambda^2-c)^2};$$

$$(12) \qquad \int\frac{dx}{(x-\rho)\sqrt{a+bx+cx^2}}=\int\frac{-2\,d\lambda}{(\alpha-\rho)\lambda^2-c(\beta-\rho)}.$$

Here, $\alpha-\rho$ and $\beta-\rho$ are both positive or both negative, and the value of the last integral is

$$\frac{-2}{\alpha-\rho}\sqrt{\frac{\alpha-\rho}{-c(\beta-\rho)}}\tan^{-1}\left(\lambda\sqrt{\frac{\alpha-\rho}{-c(\beta-\rho)}}\right).$$

We have, then, finally:

$$(13) \qquad \int\frac{dx}{(x-\rho)\sqrt{a+bx+cx^2}}$$

$$=\frac{2}{\sqrt{-f(\rho)}}\tan^{-1}\left(\frac{\rho-\alpha}{\sqrt{-f(\rho)}}\cdot\frac{\sqrt{a+bx+cx^2}}{x-\alpha}\right),$$

$$f(\rho)=a+b\rho+c\rho^2<0.$$

If $c > 0$, the conic is a hyperbola. We have:

$$a + bx + cx^2 = c(x - \alpha)(x - \beta),$$

where the parentheses are either both positive or both negative. The formal work proceeds as before, and we arrive at the integral (12). Here, however, $\alpha - \rho$ and $\beta - \rho$ have opposite signs, but c is positive. Hence the subsequent work is unchanged, and formula (13) covers this case, also.

If $\rho = \alpha$ or β, the integration can be effected directly by the above substitution. The result is algebraic.

Example 3. $$\int \frac{dx}{(x^2 + px + q)\sqrt{a + bx + cx^2}}, \qquad p^2 - 4q < 0.$$

If the roots of the radicand are real, the rationalization used in Example 2 is expedient. The computation can be carried through with elegance; but the final result is a complicated formula.

When the roots of the radicand are imaginary, the problem is still more complex. A method of treatment which applies to both cases consists in making a fractional linear transformation of x whereby the linear term in each quadratic polynomial is eliminated.* This can be done piecemeal. First reduce the radicand to the form $1 + x^2$:

$$\int \frac{dx}{(x^2 + px + q)\sqrt{1 + x^2}}.$$

Next, make the linear transformation:

$$t = \frac{x + h}{1 - hx}, \qquad x = \frac{t - h}{1 + ht},$$

where h is determined by the equation:

$$ph^2 + 2(1 - q)h - p = 0.$$

Then

$$1 + x^2 = \frac{(1 + h^2)(1 + t^2)}{(1 + ht)^2}, \qquad dx = \frac{(1 + h^2)\,dt}{(1 + ht)^2},$$

$$x^2 + px + q = \frac{A + Bt^2}{(1 + ht)^2},$$

$$A = h^2 - ph + q > 0, \qquad B = 1 + ph + qh^2 > 0;$$

$$\int \frac{dx}{(x^2 + px + q)\sqrt{1 + x^2}} = \sqrt{1 + h^2}\int \frac{\pm\,(1 + ht)\,dt}{(A + Bt^2)\sqrt{1 + t^2}}.$$

* Cf. Chap. IX, § 4.

where the lower sign is to be chosen in case $1 + ht$ is a negative quantity.

We are thus led (on replacing t by x) to the integrals:

$$\int \frac{dx}{(A + Bx^2)\sqrt{1 + x^2}}, \qquad \int \frac{x\,dx}{(A + Bx^2)\sqrt{1 + x^2}}; \qquad A > 0, \quad B > 0.$$

The integrands can be rationalized as follows. Cut the conic

$$y^2 = 1 + x^2$$

by a pencil of lines parallel to an asymptote:

(14) $$y = -x + \lambda.$$

There is no difficulty in carrying through the details of the work, but the final result is in form less simple than that obtained by a device. It is obvious that the second integral can be readily evaluated by means of the substitution: $z = \sqrt{1 + x^2}$. And now the first integral can be reduced to the second by the transformation:

$$y = 1/x.$$

The method applies to the more general case that the radicand is $C + Dx^2$, where C and D are any numbers not both negative, and neither zero. The result is as follows: *

(15) $$\int \frac{dx}{(A + Bx^2)\sqrt{C + Dx^2}} = \frac{1}{A}\sqrt{\frac{A}{BC - AD}}\tan^{-1}\left(x\sqrt{\frac{BC - AD}{A(C + Dx^2)}}\right),$$

$$\text{or } = \frac{1}{2A}\sqrt{\frac{A}{AD - BC}}\log\frac{\sqrt{C + Dx^2} + x\sqrt{(AD - BC)/A}}{\sqrt{C + Dx^2} - x\sqrt{(AD - BC)/A}}.$$

(16) $$\int \frac{x\,dx}{(A + Bx^2)\sqrt{C + Dx^2}} = \frac{1}{B}\sqrt{\frac{B}{AD - BC}}\tan^{-1}\sqrt{\frac{B(C + Dx^2)}{AD - BC}},$$

$$\text{or } = \frac{1}{2B}\sqrt{\frac{B}{BC - AD}}\log\frac{\sqrt{C + Dx^2} - \sqrt{(BC - AD)/B}}{\sqrt{C + Dx^2} + \sqrt{(BC - AD)/B}}.$$

* Formulas 229, 230 of Peirce's *Tables* are simpler than these; but the $\tan^{-1}$ formulas there tabulated are wrong except under restrictions not there stated. Thus the first formula, 229, is true when $a' > 0$; but when $a' < 0$, there should be a $-$ sign before $\tan^{-1}$. Again, no one of the formulas 229 covers the case:

$$a' < 0, \qquad a'c - ac' < 0.$$

EXERCISES

1. By means of the rationalization set forth in Example 3, (14), when $c > 0$, namely, $y = -x\sqrt{c} + \lambda$, obtain Formula 160 of the *Tables:*

$$\int \frac{dx}{\sqrt{a + bx + cx^2}} = \frac{1}{\sqrt{c}} \log\left(\sqrt{a + bx + cx^2} + x\sqrt{c} + \frac{b}{2\sqrt{c}}\right).$$

2. Use the other asymptote, and show that the result thus obtained reduces to the above.

3. Prove the following identities for the principal values of the functions entering.

(a) $$\tan^{-1} u = \sin^{-1}\frac{u}{\sqrt{1+u^2}}; \qquad -\infty < u < +\infty;$$

(b) $$2\tan^{-1} u = \tan^{-1}\frac{2u}{1-u^2}, \qquad -1 < u < 1.$$

(c) $$\tan^{-1} u = \cos^{-1}\frac{1}{\sqrt{1+u^2}}, \qquad 0 \leqq u < +\infty.$$

4. By means of the rationalization illustrated in Example 2 evaluate the integral

$$\int \frac{dx}{\sqrt{3 - 2x - x^2}}. \qquad \textit{Ans.} \quad -2\tan^{-1}\sqrt{\frac{1-x}{x+3}}.$$

5. Reduce the answer in Question 4 to the form given by the *Tables*, Formula 161.

6. Compute the integral:

$$\int \frac{dx}{(8 + 12x + 5x^2)\sqrt{5 + 6x + 2x^2}}.$$

7. Show that the substitution of § 9,

$$\lambda = \tan\frac{\theta}{2},$$

whereby $\sin\theta$ and $\cos\theta$ took on the forms

$$\sin\theta = \frac{2\lambda}{1+\lambda^2}, \qquad \cos\theta = \frac{1-\lambda^2}{1+\lambda^2},$$

amounts in substance to the rationalization of the circle

$$x^2 + y^2 = 1$$

by means of the pencil of lines through the point $(-1, 0)$, namely:

$$y = \lambda(x + 1).$$

8. *Evaluate:* $\displaystyle\int\frac{dx}{(x-\rho)\sqrt{c(x-\alpha)(x-\beta)}}$, $\rho=\alpha$ or β.

9. Give all the details of the proof of Formulas (15), (16), considering, when necessary, the case that $x<0$.

13. Conclusion; the Actual Computation. Any rational function $R(x, y)$ can be written in the form:

$$R(x, y)=\frac{g(x, y)}{G(x, y)},$$

where $g(x, y)$ and $G(x, y)$ are polynomials. If

$$y=\sqrt{a+bx+cx^2},$$

the even powers of y are polynomials in x and can be replaced by these values. The odd powers can be written each as the product of y by an even power, and the latter factor can be replaced by a polynomial. Thus R reduces to the form:

$$R(x, y)=\frac{A(x)+y\,B(x)}{C(x)+y\,D(x)},$$

where $A(x), \ldots, D(x)$ are polynomials.

The denominator can be rationalized in the usual way by multiplying numerator and denominator by $C-yD$. Thus

$$R(x, y)=\rho(x)+y\,\sigma(x),$$

where $\rho(x)$, $\sigma(x)$ are rational functions. Finally we can write:

$$R(x, y)=\rho(x)+\tau(x)\cdot\frac{1}{y},$$

where τ is rational.

Turning now to the integration of R:

$$\int R(x, y)\,dx=\int\rho(x)\,dx+\int\tau(x)\frac{dx}{y},$$

we have first the integral of a rational function, and the method of partial fractions, as above set forth, leads to the desired evaluation. In the second integral, let $\tau(x)$ be expressed in terms of partial fractions. We are thus led to integrals of the following types:

$$\int\frac{x^n\,dx}{y},\quad \int\frac{dx}{(x-\rho)^n y},\quad \int\frac{dx}{(x^2+px+q)^n y},\quad \int\frac{x\,dx}{(x^2+px+q)^n y}.$$

These integrals are computed by the aid of Reduction Formulas; cf. Chap. II.

14. Integration by Parts. The method is contained in the formula

$$(1) \qquad \int u\,dv = uv - \int v\,du;$$

Introduction to the Calculus, p. 243. The cases to which the method applies form a restricted class, but one which, in an extended study of integration, must be recognized.

The method is best studied through typical examples like those of the paragraph to which reference has been made, and the student should now review that paragraph. There is little point in multiplying examples here, since such examples would carry with them the direction to use this method, and the whole difficulty lies in the fact that, in practice, the student is *not* told when to use the method. For this reason he should strive to detect those integrals in the miscellaneous list at the end of the chapter (cf. also § 10) which are best evaluated in this manner.* Perhaps a single example may be useful.

Example. To evaluate the integral

$$I = \int x \tan^{-1}\frac{a^2 + x^2}{a^2}\,dx.$$

Let

$$u = \tan^{-1}\frac{a^2 + x^2}{a^2}, \qquad dv = x\,dx,$$

the object being to eliminate the transcendental function through differentiation. Then

$$du = \frac{2\,a^2 x\,dx}{a^4 + (a^2 + x^2)^2}, \qquad v = \frac{x^2}{2},$$

and we have by (1):

$$I = \frac{x^2}{2}\tan^{-1}\frac{a^2 + x^2}{a^2} - a^2 \int \frac{x^3\,dx}{a^4 + (a^2 + x^2)^2}.$$

To evaluate the latter integral, let $y = x^2$:

$$\int \frac{y\,dy}{a^4 + (a^2 + y)^2} = \tfrac{1}{2}\log\left[a^4 + (a^2 + y)^2\right] - \tan^{-1}\frac{a^2 + y}{a^2}.$$

* The reduction formulas of Chap. II are frequently presented as illustrations of the method; but they are spurious illustrations, since the direct manner of arriving at them is through the differential formula

$$d(uv) = u\,dv + v\,du,$$

and not through its integral form, (1), of the text.

Hence, finally,

$$\int x \tan^{-1}\frac{a^2+x^2}{a^2}\,dx = \frac{x^2+a^2}{2}\tan^{-1}\frac{a^2+x^2}{a^2} - \frac{a^2}{4}\log\left[a^4+(a^2+x^2)^2\right].$$

The student may like to evaluate

$$\int x^3 \tan^{-1}\frac{a^2+x^2}{a^2}\,dx.$$

EXERCISES ON CHAPTER I

Evaluate the following integrals, using the methods, but not the formulas, developed in the text.

1. $\displaystyle\int \frac{x\,dx}{\sqrt{a^2-x^2}}.$ 2. $\displaystyle\int \frac{(e^x-e^{-x})^2\,dx}{e^x}.$ 3. $\displaystyle\int x\cos x^2\,dx.$

4. $\displaystyle\int \frac{\sin(\pi\sqrt{s}-\delta)\,ds}{\sqrt{s}}.$ 5. $\displaystyle\int \frac{2\cos\alpha x - \sin\beta x}{\alpha+\beta}\,dx.$ 6. $\displaystyle\int \frac{e^{-\frac{1}{x}}}{x^2}\,dx.$

7. $\displaystyle\int \frac{e^{-x}\,dx}{e^x+e^{-x}}.$ 8. $\displaystyle\int \sqrt{\frac{x}{2+3x}}\,dx.$ 9. $\displaystyle\int \frac{dx}{x\log x^2}.$

10. $\displaystyle\int \frac{d\theta}{5-\sin 2\theta}.$ 11. $\displaystyle\int x\cot^2 x\,dx.$ 12. $\displaystyle\int \frac{\cos\theta\,d\theta}{\cos 2\theta}.$

13. $\displaystyle\int \frac{x\,dx}{(a^2+x^2)\sqrt{4+x^2}}.$ 14. $\displaystyle\int \log(1+x+x^2)\,dx.$

15. $\displaystyle\int \frac{d\theta}{1+\tan\theta}.$ 16. $\displaystyle\int x^2\tan^{-1}x\,dx.$ 17. $\displaystyle\int \frac{x^5\,dx}{(1-x)^4}.$

18. $\displaystyle\int \frac{dx}{(1+x^2)\tan^{-1}x}.$ 19. $\displaystyle\int \frac{d\theta}{\sin^2\theta-2\cos^2\theta}.$

20. $\displaystyle\int x\sin^{-1}\sqrt{\frac{x}{1+x}}\,dx.$ 21. $\displaystyle\int e^{-x^2}(2x-3x^3)\,dx.$

22. $\displaystyle\int \cos^5\theta\,d\theta.$ 23. $\displaystyle\int x\tan\pi x^2\,dx.$ 24. $\displaystyle\int \frac{\log x\,dx}{x}.$

25. $\displaystyle\int \frac{\sin\theta-\cos\theta}{1+\sin\theta+\cos\theta}\,d\theta.$ 26. $\displaystyle\int \sin px\cos qx\,dx.$

28. $\displaystyle\int \frac{(1+a\cos\theta)\,d\theta}{(1+a^2)\cos^2\theta+1}.$ 29. $\displaystyle\int (a+x)\log(x+\sqrt{a^2-x^2})\,dx.$

30. $\displaystyle\int \sin m\theta\sin n\theta\,d\theta.$ 31. $\displaystyle\int \frac{\sqrt{a^2+b^2+a^2\sin^2\theta}}{\cos\theta}\,d\theta.$

32. $\int \frac{x + \sin x}{1 + \cos x}\,dx.$ **33.** $\int \frac{d\theta}{\sin^3 \frac{1}{2}\theta}.$ **34.** $\int e^{-x} \sin^2 x\,dx.$

35. $\int \csc^{-1} x\,dx.$ **36.** $\int x\,e^{-\sqrt{x}}\,dx.$ **37.** $\int \frac{dx}{1 + 7\cos x}.$

38. $\int \frac{1 + \tan x}{\cos^2 x}\,dx.$ **39.** $\int \frac{\tan x\,dx}{1 - \tan^2 x}.$ **40.** $\int \frac{\sin x\,dx}{\sin (a + x)}.$

41. $\int (a^2 + x^2) \log (a^2 + x^2)\,dx.$ **42.** $\int \frac{(r\cos\theta - a)\,dr}{[a^2 - 2\,ar\cos\theta + r^2]^{\frac{3}{2}}}.$

43. $\int \frac{(a^2 - r^2)\,d\theta}{a^2 - 2\,ar\cos\theta + r^2}.$ **44.** $\int e^{-t} \sin (nt - \gamma)\,dt.$

45. $\int \frac{x^5 + a^5}{x^3 + a^3}\,dx.$ **46.** $\int \frac{\log\sqrt{1 + x^2}\,dx}{(1 - x)^2}.$ **47.** $\int \frac{dx}{x\sqrt{1 + x^2}}.$

49. $\int x\,e^{ax} \cos bx\,dx.$ **50.** $\int \frac{dx}{\sqrt{x} + \sqrt{1 - x}}.$

51. $\int \frac{\sqrt{1 + x} + \sqrt{1 - x}}{\sqrt{1 + x} - \sqrt{1 - x}}\,dx.$ **52.** $\int \frac{dx}{(x - 2)\sqrt{x^2 - 4\,x + 3}}.$

53. Compute the value of the definite integral

$$\int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} \frac{dx}{5 + 3\cos x}.$$ *Ans.* 1.1071.

54 Compute the value of the definite integral

$$\int_{\frac{3\pi}{2}}^{2\pi} \frac{dx}{3 + 5\cos x}.$$ *Ans.* 0.27465.

CHAPTER II

REDUCTION FORMULAS

1. The Integral $\int \sin^n x \cos^m x \, dx$.

We begin by taking the differential of a function of the same type as the integrand:

(1) $d(\sin^\nu x \cos^\mu x) = \nu \sin^{\nu-1} x \cos^{\mu+1} x \, dx - \mu \sin^{\nu+1} x \cos^{\mu-1} x \, dx.$

If we integrate each side of this equation, we obtain a relation between the integrals

$$\int \sin^{\nu-1} x \cos^{\mu+1} x \, dx, \qquad \int \sin^{\nu+1} x \cos^{\mu-1} x \, dx.$$

Thus if we wished, in the given integral, to increase n and decrease m (or vice versa), we could effect the result. But this is a very special and relatively unimportant case. It does not enable us to change one of the exponents without changing the other. We are led, therefore, to make a trigonometric reduction. Write

$$\cos^{\mu+1} x = \cos^{\mu-1} x \cos^2 x = \cos^{\mu-1} x - \cos^{\mu-1} x \sin^2 x.$$

Then equation (1) becomes:

(2) $d(\sin^\nu x \cos^\mu x) = \nu \sin^{\nu-1} x \cos^{\mu-1} x \, dx - (\mu + \nu) \sin^{\nu+1} x \cos^{\mu-1} x \, dx.$

On integrating this equation, we obtain a formula whereby the exponent of the cosine factor is unchanged, but the exponent of the sine factor is changed by 2. Let

$$\nu + 1 = n, \qquad \mu - 1 = m.$$

Then

$$\sin^{n-1} x \cos^{m+1} x = (n-1) \int \sin^{n-2} x \cos^m x \, dx - (n+m) \int \sin^n x \cos^m x \, dx,$$

(3) $$\int \sin^n x \cos^m x \, dx = -\frac{\sin^{n-1} x \cos^{m+1} x}{n+m} + \frac{n-1}{n+m} \int \sin^{n-2} x \cos^m x \, dx.$$

In particular we obtain on setting $m = 0$:

(4) $$\int \sin^n x\,dx = -\frac{\sin^{n-1} x \cos x}{n} + \frac{n-1}{n}\int \sin^{n-2} x\,dx.$$

Thus for $n = 2, 4$ we have:

$$\int \sin^2 x\,dx = -\tfrac{1}{2}\sin x \cos x + \tfrac{1}{2}x;$$

$$\int \sin^4 x\,dx = -\tfrac{1}{4}\sin^3 x \cos x + \tfrac{3}{4}\int \sin^2 x\,dx$$

$$= -\tfrac{1}{4}\sin^3 x \cos x - \tfrac{3}{8}\sin x \cos x + \tfrac{3}{8}x,$$

and by setting $n = 6$ the student can now verify the result of Ex. 5 in Chap. IX, § 8, of the *Introduction to the Calculus*.

We must warn the student, however, against the stupidity of applying any of these reduction formulas when there is an obvious short cut. Thus

$$\int \sin^3 x\,dx = -\int (1 - \cos^2 x)\,d\cos x = -\cos x + \tfrac{1}{3}\cos^3 x,$$

and to use Formula (4), $n = 3$, to evaluate this integral would be much like multiplying 700 and 800 together by logarithms.

If n is negative, we wish to increase it, and so Formulas (3) and (4) should be used backwards, *i.e.* solved for the integral on the right-hand side. A neater form for the result is obtained by going back to (2) and setting

$$\nu - 1 = -n, \qquad \mu - 1 = m:$$

(5) $$\int \frac{\cos^m x\,dx}{\sin^n x} = -\frac{\cos^{m+1} x}{(n-1)\sin^{n-1} x} + \frac{n-m-2}{n-1}\int \frac{\cos^m x\,dx}{\sin^{n-2} x}.$$

On setting $m = 0$, we have:

(6) $$\int \frac{dx}{\sin^n x} = -\frac{\cos x}{(n-1)\sin^{n-1} x} + \frac{n-2}{n-1}\int \frac{dx}{\sin^{n-2} x}.$$

If in (5) n and m are equal, there is a simpler formula. Here,

$$\frac{\cos^n x}{\sin^n x} = \cot^n x = \cot^{n-2} x(\csc^2 x - 1).$$

Hence

(7) $$\int \cot^n x\,dx = -\frac{\cot^{n-1} x}{n-1} - \int \cot^{n-2} x\,dx.$$

In all these formulas, n and m may be fractional or incommensurable. But in that case the given integral cannot in general be evaluated, and so the formulas are important chiefly in the case that n and m are integers.

Formula (2) and the corresponding integral relation are always true. In passing, however, to the later formulas, a division has taken place, and it is tacitly assumed that the divisor is not zero. If it were, the resulting formula would have no meaning. Thus no danger can arise, for a formula that has no meaning cannot lead to a wrong result. Whenever one of these formulas has a meaning, it is correct.

EXERCISES

Obtain the following reduction formulas.

1. $$\int \sin^n x \cos^m x\, dx = \frac{\sin^{n+1} x \cos^{m-1} x}{m+n} + \frac{m-1}{m+n}\int \sin^n x \cos^{m-2} x\, dx.$$

2. $$\int \cos^m x\, dx = \frac{\sin x \cos^{m-1} x}{m} + \frac{m-1}{m}\int \cos^{m-2} x\, dx.$$

3. $$\int \frac{\sin^n x\, dx}{\cos^m x} = \frac{\sin^{n+1} x}{(m-1)\cos^{m-1} x} + \frac{m-n-2}{m-1}\int \frac{\sin^n x\, dx}{\cos^{m-2} x}.$$

4. $$\int \frac{dx}{\cos^m x} = \frac{\sin x}{(m-1)\cos^{m-1} x} + \frac{m-2}{m-1}\int \frac{dx}{\cos^{m-2} x}.$$

5. $$\int \tan^n x\, dx = \frac{\tan^{n-1} x}{n-1} - \int \tan^{n-2} x\, dx.$$

6. Obtain the formula of Question 2 directly by starting with

$$d(\sin x \cos^\mu x).$$

7. Obtain the formula of Question 4 in a similar manner.

8. Check the formulas of the Exercises against the corresponding formulas of the text by setting $x = \frac{1}{2}\pi - y$.

9. Obtain the formula of Question 5 by starting with $d \tan^\mu x$ and making a suitable trigonometric reduction of the result.

10. Evaluate the following integrals:

(a) $\int \cos^4 x\, dx.$ (b) $\int \cos^3 x\, dx.$ (c) $\int \sin x \cos^3 x\, dx.$

(d) $\int \frac{dx}{\cos^4 x}.$ (e) $\int \frac{dx}{\cos^3 x}.$ (f) $\int \frac{\sin x\, dx}{\cos^3 x}.$

2. The Integral $\int \frac{dx}{(a^2+x^2)^n}$.

It would be a false analogy with the example of the preceding paragraph to start with $d(a^2+x^2)^{-n}$, since the result would be but a single term, and that not yielding an integral of useful type. We need a *product*, and Exercises 6 and 7 in § 1 suggest the plan: *

$$(1)\qquad d[x(a^2+x^2)^{-m}] = (a^2+x^2)^{-m}\,dx - 2\,mx^2(a^2+x^2)^{-m-1}\,dx.$$

In the last term, write the factor x^2 in the form:

$$x^2 = (a^2+x^2) - a^2.$$

Thus

$$(2)\quad d[x(a^2+x^2)^{-m}] = (1-2\,m)(a^2+x^2)^{-m}\,dx + 2\,m\,a^2(a^2+x^2)^{-m-1}\,dx.$$

It is now clear that, on setting $m+1=n$ and integrating, we shall have a reduction formula worth while, namely:

$$(3)\quad \int \frac{dx}{(a^2+x^2)^n} = \frac{x}{2(n-1)a^2(a^2+x^2)^{n-1}} + \frac{2n-3}{(2n-2)a^2}\int \frac{dx}{(a^2+x^2)^{n-1}}.$$

It is this formula which occupies a pivotal position in the proof that every rational function can be integrated.

EXERCISE

In Formula (3) set $x = a\tan\theta$. Hence show how (3) can be deduced from the formula of Exercise 2, § 1.

3. The Integral $\int \frac{x^m\,dx}{\sqrt{a+bx+cx^2}}$.

Let
$$y = \sqrt{a+bx+cx^2}$$
and take $d(x^m y)$:

$$(1)\qquad d(x^m y) = mx^{m-1}y\,dx + \frac{(\tfrac{1}{2}b+cx)x^m}{y}\,dx,$$

$$mx^{m-1}y = \frac{mx^{m-1}(a+bx+cx^2)}{y}.$$

* There is no inductive treatment possible in this part of integration. Imagination, resourcefulness, and the power of keen observation are the qualities required.

Hence

(2) $$d(x^m y) = ma\frac{x^{m-1}}{y}\,dx + (m+\tfrac{1}{2})b\frac{x^m}{y}\,dx + (m+1)c\frac{x^{m+1}}{y}\,dx.$$

Let $m+1=n$, and integrate. Thus we obtain the reduction formula:

(3) $$\int\frac{x^n\,dx}{y} = \frac{x^{n-1}y}{nc} - \frac{2n-1}{2n}\frac{b}{c}\int\frac{x^{n-1}\,dx}{y} - \frac{n-1}{n}\frac{a}{c}\int\frac{x^{n-2}\,dx}{y}.$$

By taking $d(y/(x-\rho)^m)$ and proceeding in a similar manner we obtain the reduction formulas

(4) $$\int\frac{dx}{(x-\rho)^n y} = -\frac{y}{(n-1)f(\rho)(x-\rho)^{n-1}} - \frac{2n-3}{2n-2}\frac{f'(\rho)}{f(\rho)}\int\frac{dx}{(x-\rho)^{n-1}y} - \frac{n-2}{n-1}\frac{c}{f(\rho)}\int\frac{dx}{(x-\rho)^{n-2}y},$$

$$f(\rho) = a + b\rho + c\rho^2 \neq 0.$$

(5) $$\int\frac{dx}{(x-\rho)^n y} = -\frac{y}{(n-\frac{1}{2})f'(\rho)(x-\rho)^n} - \frac{2n-2}{2n-1}\frac{c}{f'(\rho)}\int\frac{dx}{(x-\rho)^{n-1}y},$$

$$f(\rho) = 0, \qquad f'(\rho) \neq 0.$$

Reduction formulas for the integrals

(6) $$\int\frac{dx}{(x^2+px+q)^n y}, \qquad \int\frac{x\,dx}{(x^2+px+q)^n y}, \quad p^2 - 4q < 0,$$

can be obtained by considering simultaneously

$$d[y(x^2+px+q)^{-m}] \quad\text{and}\quad d[xy(x^2+px+q)^{-m}].$$

EXERCISES

1. Obtain a reduction formula for

$$\int\frac{x^n\,dx}{\sqrt{a+bx}}.$$

2. Give the details of obtaining the reduction formulas (4) and (5).

3. Obtain a reduction formula for

$$\int\frac{dx}{(x^2+b^2)^n\sqrt{a^2+x^2}}.$$

4. Obtain reduction formulas for the integrals (6).

5. Starting with $d\{(\lambda x + \mu)X^{-m}\}$, where $X = a + bx + cx^2$ and λ, μ are undetermined constants, show how, by a suitable choice of λ, μ, reduction formulas may be obtained for the integrals

$$\int \frac{dx}{X^n}, \qquad \int \frac{x\,dx}{X^n}.$$

6. Develop reduction formulas for

$$\int x^n \sin x\,dx \qquad \text{and} \qquad \int x^n \cos x\,dx.$$

CHAPTER III

DOUBLE INTEGRALS

1. The Double Integral.* The definite integral of a function of a single variable,

$$\int_a^b f(x)\,dx,$$

was defined as follows; *Introduction to the Calculus*, Chap. 12, § 3. The interval $a \leqq x \leqq b$ was divided up into n parts by the points $x_0 = a, x_1, \cdots, x_{n-1}, x_n = b$. The function $f(x)$, which was required to be continuous throughout this interval, was formed for each of the points $x_1', x_2', \cdots, x_n'$, where x_k' is an arbitrary point of the k-th subinterval,

$$x_{k-1} \leqq x_k' \leqq x_k,$$

and the products,

$$f(x_k')\,\Delta x_k, \qquad \Delta x_k = x_k - x_{k-1},$$

were added:

$$\sum_{k=1}^{n} f(x_k')\,\Delta x_k = f(x_1')\,\Delta x_1 + f(x_2')\,\Delta x_2 + \cdots + f(x_n')\,\Delta x_n.$$

The limit approached by this sum as n increases without limit, the longest Δx_k approaching 0, was defined as the *definite integral of* $f(x)$ *from* a *to* b*:*

* The student who is approaching this subject for the first time should study carefully §§ 1–5. He may then choose between §§ 6, 7, 8, for an intensive study of any one of these three paragraphs will serve the present purpose.

He should next study §§ 9, 10 with attention to every detail, and master thoroughly § 10. He can then take §§ 11, 12, 13 in any order. It is more important that he do thoroughly one or two of these paragraphs than that he cover all three superficially.

He will do well to leave the chapter at this point and turn to the next chapter, that on Triple Integrals. The treatment here brings out strongly the geometric side of the subject, and is not encumbered by examples which present analytical difficulties. But the geometry involved in determining the limits of integration he *must* master. He can then turn back to his double integrals and, after a careful review of §§ 3 and 10, take the remainder of the chapter in any order.

$$\lim_{n=\infty} \sum_{k=1}^{n} f(x'_k)\,\Delta x_k = \int_a^b f(x)\,dx. \tag{1}$$

The foregoing conception, or definition, can be extended to functions of two variables as follows. Let $f(x, y)$ be a continuous function of the two independent variables x, y throughout a finite region S of the (x, y)-plane. Let S be divided in any manner into n sub-regions, of area ΔS_1, ΔS_2, $\cdots$, ΔS_n, which together just fill out the region S. Let (x_k, y_k) be a point chosen arbitrarily within or on the boundary of the k-th region. Form the sum

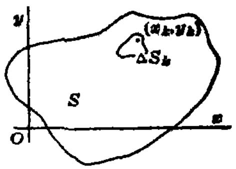


FIG. 2

$$\sum_{k=1}^{n} f(x_k, y_k)\Delta S_k = f(x_1, y_1)\Delta S_1 + f(x_2, y_2)\Delta S_2 + \cdots + f(x_n, y_n)\Delta S_n. \tag{2}$$

The limit approached by this sum as n increases without limit, the greatest diameter of any ΔS_k approaching 0, is defined as the *double integral of $f(x, y)$, extended over the region S:*

$$\lim_{n=\infty} \sum_{k=1}^{n} f(x_k, y_k)\,\Delta S_k = \iint_S f(x, y)\,dS. \tag{3}$$

The *form* of the sub-regions is wholly arbitrary. Thus they may be taken as rectangles (cf. Fig. 23, § 17); or as the four-sided figures of Fig. 24, § 18. In these cases, there will in general be irregular sub-regions along the boundary, consisting of pieces of the rectangles, etc. But it is easily seen (cf. § 2) that the contribution to the sum arising from all such regions is small, and approaches 0 as its limit. Hence such regions may be wholly suppressed; or they may be replaced each by the complete rectangle, etc.; or some may be suppressed and others completed. The value of the limit of the sum will be the same in all cases.

EXERCISE

Extend the conception to space, starting out with a three-dimensional region, V, and a function, $f(x, y, z)$, continuous throughout V. The result is the triple integral,

$$\iiint_V f(x, y, z)\,dV.$$

Give all the details of the definition.

2. Geometrical Interpretation. Law of the Mean. If we plot the surface

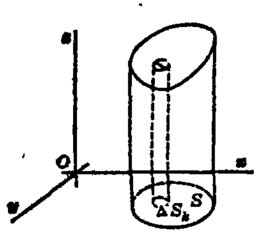


FIG. 3

(4) $$z=f(x, y)$$

and consider the cylinder standing on the region S, the volume, V, of so much of this cylinder as is cut off by the (x, y)-plane and the surface (4) represents precisely the double integral:

(5) $$V=\int_S\int f(x, y)\, dS.$$

For, each term $f(x_k, y_k)\Delta S_k$ in the sum (2) is the volume of the slender cylinder which stands on the k-th sub-region and reaches up to a point of the surface (4) lying in this cylinder. Obviously, the sum of all these n volumes differs but slightly from the volume V; and the discrepancy grows smaller and smaller and approaches 0 as its limit, when n increases without limit.*

Thus the analogue of the definite integral (1), interpreted as the *area under the curve* $y=f(x)$, is here the definite integral (3), interpreted as the *volume under the surface* $z=f(x, y)$.

We have tacitly assumed that $f(x, y)$ is positive. If $f(x, y)$ is negative, the column lies below the (x, y)-plane, and the double integral is still numerically equal to the volume of the column, but is now seen to be negative. If, finally, $f(x, y)$ is positive in some parts of S and negative in others, the double integral is seen to be equal to the algebraic sum of the volumes of those parts of the column which lie above the (x, y)-plane, taken positively, and those parts which lie below this plane, taken negatively.†

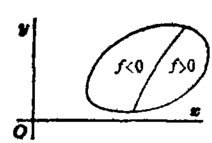


FIG. 4

Law of the Mean. It is clear that the volume, V, of the column is the same as that of a cylindrical column of like cross-section and of altitude $\bar{z}$ intermediate between the smallest and the largest values of $f(x, y)$ in S. The function f takes on this value in any one of an

* A suggestive model of the whole set of slender columns is found in the mineralogical specimens of stibnite and tourmaline.

† We are tacitly restricting ourselves to the case — sufficiently general for the needs of practice — that the number of parts in each class is finite. Otherwise, a limiting process would be necessary.

infinite number of suitably chosen points, (ξ, η), in S, and thus

$$(6) \qquad \iint_S f(x, y)\, dS = f(\xi, \eta)\, A,$$

where A denotes the area of S. This theorem is known as the *Law of the Mean*.

3. Computation of the Volume V by the Iterated Integral. We have learned how to compute all sorts of irregular volumes by the method of slicing and the application of the definite integral of a function of a single variable.* The method may be formulated generally for a solid of any shape. Assume a line in space, whose direction is taken at pleasure, and cut the solid by a variable plane perpendicular to this line; cf. Fig. 5. Denote the distance of an arbitrary point on the line from a fixed point of the line by x. The area of the cross-section made by the above plane is a function of x, which we will denote by $A(x)$, or simply A. Let the minimum x corresponding to one of the above planes be $x = a$, the maximum, $x = b$. Divide the interval from a to b into n equal parts by the points $x_0 = a, x_1, \cdots, x_n = b$ and pass planes through these points perpendicular to the line. Then the volume in question is given approximately by the sum:

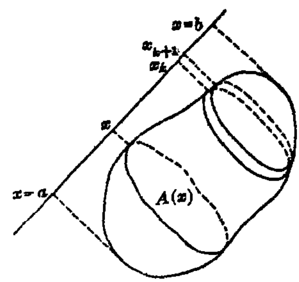


FIG. 5

$$A(x_1)\Delta x + A(x_2)\Delta x + \cdots + A(x_n)\Delta x,$$

and the limit of this sum, when n becomes infinite, is exactly the volume sought:

$$(1) \qquad V = \int_a^b A\, dx.$$

Example. To compute the volume of the ellipsoid:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1.$$

Here the cross-section made by an arbitrary plane $x = x'$ is the ellipse

$$\frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 - \frac{x'^2}{a^2}.$$

* The student will do well to work again at this stage a number of the problems on volumes in the *Introduction to the Calculus*, pp. 319–323.

Its semiaxes have the lengths $b\sqrt{1-\frac{x'^2}{a^2}}, \quad c\sqrt{1-\frac{x'^2}{a^2}}$,
and hence its area is, the accents being suppressed:

$$A = \pi bc\left(1-\frac{x^2}{a^2}\right).$$

The volume V is, therefore,

$$V = \pi bc\int_{-a}^{a}\left(1-\frac{x^2}{a^2}\right)dx = \pi bc\left(x-\frac{x^3}{3a^2}\right)\Bigg|_{-a}^{a} = \tfrac{4}{3}\pi abc.$$

It is possible to check this formula by setting $a=b=c$. We then have a sphere of radius $R=a$, and its volume is $V=\frac{4}{3}\pi a^3$.

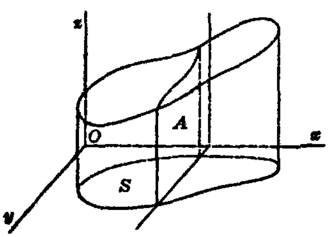


Fig. 6

The Volume, V, of § 2. To compute this volume, we cut the column by an arbitrary plane, $x=x'$, and determine the area, A, of this cross-section. Now A is simply the area under the curve

$$z=f(x', y) \qquad (x', \text{ constant})$$

between the ordinates corresponding to the abscissas $y=Y_0$ and $y=Y_1$.

Hence

$$A=\int_{Y_0}^{Y_1} f(x', y)\,dy.$$

Dropping the accent, which has now served its purpose, we have:

$$(2) \qquad A(x)=\int_{Y_0}^{Y_1} f(x, y)\,dy,$$

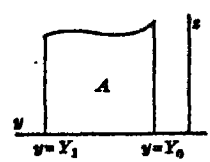


Fig. 7

where we must remember that x is constant, y being the variable of integration. The limits of integration, Y_0 and Y_1, are functions of x. If the equation of the lower boundary of S be written in the form:

$$y=\phi_0(x),$$

then $Y_0=\phi_0(x)$. And similarly, if

$$y=\phi_1(x)$$

be the equation of the upper boundary, then $Y_1=\phi_1(x)$.

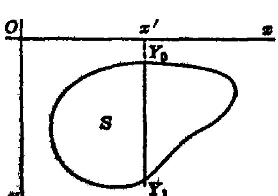


Fig. 8

It remains only to integrate A with respect to x between the limits $x=a$ and $x=b$, where a is the smallest abscissa that any point in S has, and b is the largest. We thus obtain:

$$V=\int_a^b A(x)\,dx.$$

This last integral is commonly written in either of the forms: *

(3) $$\int_a^b dx \int_{Y_0}^{Y_1} f(x, y)\,dy \qquad \text{or} \qquad \int_a^b\int_{Y_0}^{Y_1} f(x, y)\,dy\,dx.$$

It is called the *iterated integral* of $f(x, y)$ (not the double integral; the latter has been explained in § 1), since it is the result of two ordinary integrations performed in succession.

Reversal of the Order of the Integrations. Instead of integrating first with regard to y and then with regard to x, we might have reversed the order, integrating first with regard to x. We should thus obtain the formula:

(4) $$V=\int_\alpha^\beta dy \int_{X_0}^{X_1} f(x, y)\,dx.$$

The student should reproduce Fig. 6, except for the intersection of the plane $x=x'$ with the solid. He should then draw the intersection of the plane $y=y'$, and formulate the area of the cross-section as an integral, constructing the figures which correspond to Figs. 7, 8. Finally, he should make clear to himself how Formula (1) applies in the present case, y here playing the rôle of the x of that formula. Thus equation (4) results.

It may happen that the boundary of S is cut by some parallels to the axis of y in more than two points, as in Fig. 9. In that case,

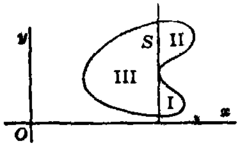


FIG. 9

* Another form sometimes employed is to be avoided, namely:

$$\int_a^b\int_{Y_0}^{Y_1} f(x, y)\,dx\,dy.$$

The second form given in the text is to be thought of as an abbreviation for

$$\int_a^b \left\{ \int_{Y_0}^{Y_1} f(x, y)\,dy \right\} dx.$$

let* S be cut up into a number of regions, each of which is of the type considered in the evaluation of V by means of the iterated integral (3). Then the portion of V which stands on any one of these regions can be evaluated by (3) and the results added.

4. The Fundamental Theorem of the Integral Calculus. The definition of the double integral, § 1:

$$\int_S\int f(x, y)\,dS,$$

is analytic, *i.e.* it is numerical, as opposed to geometrical. The infinitesimals $f(x_k, y_k)\,\Delta S_k$ are *numbers*, and the limit of their sum is a *number*.

Likewise, the iterated integral,

$$\int_a^b dx \int_{Y_0}^{Y_1} f(x, y)\,dy,$$

is analytic; it is a number.

Each of these numbers is equal to the number which is the measure of the volume, V, bounded by the surface of § 2:

$$z = f(x, y).$$

Hence these numbers are equal to each other, and thus we obtain, by the aid of geometry, a theorem of analysis, whereby an important limit is evaluated.† It may be stated as follows.

FUNDAMENTAL THEOREM OF THE INTEGRAL CALCULUS. *Let $f(x, y)$ be a continuous function of x and y throughout a region S of the (x, y)-plane. Divide this region up into n pieces of area ΔS_1, ΔS_2, $\cdots$, ΔS_n and form the sum:*

$$f(x_1, y_1)\,\Delta S_1 + f(x_2, y_2)\,\Delta S_2 + \cdots + f(x_n, y_n)\,\Delta S_n,$$

where (x_k, y_k) is any point of the k-th sub-region. If n now be allowed to increase without limit, the maximum diameter of any sub-region approaching 0 as its limit, this sum will approach a limit, namely, the double integral

$$\int_S\int f(x, y)\,dS;$$

* More precisely, we restrict ourselves to regions S which have this property.

† Compare the corresponding procedure in the *Introduction to the Calculus*, Chap. 12, § 3, whereby two expressions for the area under a curve were equated to each other.

and the value of the limit is given by the iterated integral,

$$\int_a^b dx \int_{Y_0}^{Y_1} f(x, y)\, dy \qquad or \qquad \int_\alpha^\beta dy \int_{X_0}^{X_1} f(x, y)\, dx,$$

where the limits of integration are determined as set forth in § 3.

Expressed as a formula the theorem is as follows:

$$\int_S\int f(x, y)\, dS = \int_a^b dx \int_{Y_0}^{Y_1} f(x, y)\, dy = \int_\alpha^\beta dy \int_{X_0}^{X_1} f(x, y)\, dx.$$

The abbreviated notation

$$\int\int f(x, y)\, dx\, dy$$

may mean either the double integral or the iterated integral. This notation should be used only when it is explicitly stated, or when it is clear from the context, which integral is meant.

The rôle which the Fundamental Theorem plays in the applications of the calculus is the following. The physical problems of determining masses, centres of gravity, moments of inertia, fluid pressures, attractions, etc., lead each time to a *formulation* which involves a double integral. The *computation* of the double integral can be performed by aid of the Fundamental Theorem.

5. Volumes by Double Integration. We have *formulated* the volume of a column standing on a plane region S and capped by the surface $z = f(x, y)$ by means of the double integral:

(1) $$V = \int_S\int z\, dS.$$

And we have learned how to *compute* this volume by means of the iterated integral:

(2) $$V = \int_a^b dx \int_{Y_0}^{Y_1} z\, dy.$$

The present paragraph is devoted to examples illustrating the method.

Let it be required to compute the volume cut off from the paraboloid:

(3) $$z = 1 - \frac{x^2}{4} - \frac{y^2}{9}$$

by the (x, y)-plane. Since the surface is obviously symmetric with respect both to the (x, z) and the (y, z)-plane, it is sufficient to compute the part of the volume that lies in the first octant, and then multiply the result by 4. This volume, V, is expressed by the double integral (1), extended over the surface S cut out of the first quadrant in the (x, y)-plane by the ellipse

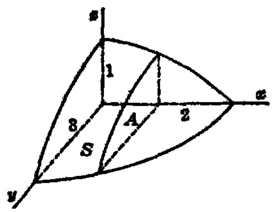


FIG. 10

$$0 = 1 - \frac{x^2}{4} - \frac{y^2}{9}. \tag{4}$$

To evaluate this double integral by means of the iterated integral (2), we cut the region S by the line $x = x'$ and consider z as a function of y alone along this line. Thus

$$\int_{Y_0}^{Y_1} z\,dy = \int_0^Y \left(1 - \frac{x'^2}{4} - \frac{y^2}{9}\right) dy,$$

where $Y = Y_1$ is the largest value y can have along the segment in question. This value is the positive ordinate of the ellipse (4) corresponding to $x = x'$:

$$0 = 1 - \frac{x'^2}{4} - \frac{y^2}{9}, \qquad Y = \tfrac{3}{2}\sqrt{4 - x'^2}.$$

Thus

$$\int_0^Y \left(1 - \frac{x'^2}{4} - \frac{y^2}{9}\right) dy = \left(1 - \frac{x'^2}{4}\right) y - \frac{y^3}{27}\bigg|_0^Y = \left\{1 - \frac{x'^2}{4} - \frac{Y^2}{27}\right\} Y$$

$$= \tfrac{1}{4}(4 - x'^2)^{\frac{3}{2}}.$$

Hence, dropping the accent, we have:

$$\int_0^Y \left(1 - \frac{x^2}{4} - \frac{y^2}{9}\right) dy = \tfrac{1}{4}(4 - x^2)^{\frac{3}{2}}.$$

The second integration, with respect to x, is from the smallest x of any point in S, here 0, to the largest x, here 2. From the *Tables*, No. 137, we have:

$$\tfrac{1}{4}\int_0^2 (4 - x^2)^{\frac{3}{2}} dx = \frac{1}{16}\left[x(4 - x^2)^{\frac{3}{2}} + 6\,x\sqrt{4 - x^2} + 24 \sin^{-1}\frac{x}{2}\right]_0^2 = \tfrac{3}{4}\pi,$$

and so the total volume is $3\pi = 9.4248$.

EXERCISES

1. Work the problem of the text by means of the other form of the iterated integral, § 3, (4).

2. If the region S is a right triangle with its vertices at (0, 0), (3, 0), (3, 2), and if the surface $z = f(x, y)$ is a plane whose intercepts on the axes are each 9, show that $V = 19$.

3. A round hole of radius unity is bored through the solid of Fig. 10, the axis of the hole being the axis of z. Find the volume removed.

4. Compute the volume of a cylindrical column standing on the area common to the two parabolas

$$x = y^2, \qquad y = x^2$$

as base and cut off by the surface

$$z = 12 + y - x^2.$$

5. Work Example 3, integrating in the other order.

6. The same for Example 4.

6. Mass of a Lamina of Variable Density.* Consider a plane lamina** of variable density, ρ. We assume that ρ is a continuous function of (x, y):

$$(1) \qquad \rho = f(x, y).$$

To find the mass of the lamina, divide the latter into n pieces and denote the area of the k-th piece by ΔS_k. Then the mass of this piece will be given by the equation:

$$(2) \qquad \Delta M_k = \bar{\rho}_k \Delta S_k,$$

where $\bar{\rho}_k$ denotes the *average density* for the piece. Moreover,

$$(3) \qquad \rho'_k \leqq \bar{\rho}_k \leqq \rho''_k,$$

where ρ'_k and ρ''_k denote respectively the least and the greatest values of ρ in the sub-region. The function $f(x, y)$, being continuous, takes on any intermediate value, as $\bar{\rho}_k$, in a suitably chosen point (x_k, y_k) of the sub-region.*** Hence

* Cf. foot-note, § 1, p. 44.

** The conception of a *lamina* and the relations (2) and (3) below are set forth in the *Introduction to the Calculus*, Chap. XII, § 10.

*** In general, there will be a curve of such points (x_k, y_k). If, in particular, ρ is constant, then ρ'_k, $\bar{\rho}_k$, and ρ''_k all coincide, and any point whatever of the sub-region may be taken as (x_k, y_k).

$$M = \sum_{k=1}^{n} \Delta M_k = \sum_{k=1}^{n} f(x_k, y_k)\,\Delta S_k. \tag{4}$$

Now let n increase without limit, the maximum diameter of any sub-region approaching 0 as its limit. Then, by definition,

$$\lim_{n=\infty} \sum_{k=1}^{n} f(x_k, y_k)\,\Delta S_k = \int\!\!\int_S f(x, y)\,dS.$$

Hence we have the mass of the lamina expressed, or formulated, as a double integral:

$$M = \int\!\!\int_S \rho\,dS. \tag{5}$$

If, in particular, the density is constant, then

$$M = \rho A,$$

where A denotes the area of the lamina.

Example. To find the mass of a square lamina whose density is proportional to the square of the distance from one corner.

Let the origin of coordinates be taken at that corner, and let the square lie in the first quadrant. Then

$$\rho = c(x^2 + y^2),$$

and

$$M = c\int\!\!\int_S (x^2 + y^2)\,dS.$$

The double integral is evaluated by means of the iterated integral:

$$\int\!\!\int_S (x^2 + y^2)\,dS = \int_0^a dy \int_0^a (x^2 + y^2)\,dx = \tfrac{2}{3}\,a^4.$$

Hence

$$M = \tfrac{2}{3}\,ca^4.$$

EXERCISES

Find the mass of each of the following laminas:

1. A right triangle, whose density is proportional to the square of the distance from an acute angle. *Ans.* $\frac{1}{12}\,c\,ab(a^2 + 3\,b^2)$.

2. The same, if the density is proportional to the square of the distance from the right angle. *Ans.* $\frac{1}{12}\,c\,ab\,(a^2 + b^2)$.

3. A semicircle, whose density is proportional to the distance from the bounding diameter.

4. A right triangle, whose density is proportional to the distance from one leg.

5. The same if the density is proportional to the distance from the hypotenuse.

6. The segment of a parabola cut off by the latus rectum, if the density is proportional to the square of the distance from the focus.

7. The same, if the density is proportional to the distance from the latus rectum.

8. At what distance from the origin are the points of the square treated in the example of the text, at which the density is equal to the average density of the lamina? *Ans.* $a\sqrt{\tfrac{2}{3}}$.

7. Centre of Gravity of a Lamina. To find the abscissa, $\bar{x}$, of the centre of gravity, G, of a lamina, divide the latter as before into n pieces and concentrate the mass of each piece at its centre of gravity, whose coordinates we will denote by (x_k, y_k). Then, by § 6 and the *Introduction to the Calculus*, Chap. XII, § 6,

$$\bar{x} = \frac{x_1\bar{\rho}_1\,\Delta S_1 + x_2\bar{\rho}_2\,\Delta S_2 + \cdots + x_n\bar{\rho}_n\,\Delta S_n}{M}.$$

If, in particular, the density is constant, $M = \rho A$, and

$$\bar{x} = \frac{x_1\,\Delta S_1 + x_2\,\Delta S_2 + \cdots + x_n\,\Delta S_n}{A}.$$

Let n increase without limit, the maximum diameter of any subregion approaching zero as its limit. Then

$$\lim_{n=\infty}\,[x_1\,\Delta S_1 + x_2\,\Delta S_2 + \cdots + x_n\,\Delta S_n] = \iint_S x\,dS.$$

Since $\bar{x}$ and A do not change with n, we have:

(1) $$\bar{x} = \frac{\displaystyle\iint_S x\,dS}{A}.$$

In this case G is often spoken of as the *centre of gravity of the plane area*, S.

If ρ is variable, but continuous, we can still proceed as before; but

(2) $$\lim_{n=\infty}[\bar\rho_1 x_1 \Delta S_1 + \bar\rho_2 x_2 \Delta S_2 + \cdots + \bar\rho_n x_n \Delta S_n]$$

is no longer by definition a double integral, since the points of the k-th sub-region, in which ρ takes on the value $\bar\rho_k$, are not seen to contain any one whose abscissa is equal to that of the centre of gravity of this sub-region, namely, x_k. Nevertheless, this limit suggests the definite integral

(3) $$\lim_{n=\infty}[\rho_1 x_1 \Delta S_1 + \rho_2 x_2 \Delta S_2 + \cdots + \rho_n x_n \Delta S_n] = \iint_S \rho\, x\, dS,$$

where ρ_k is the value of ρ in the centre of gravity (x_k, y_k) of the k-th sub-region.

That the limits (2) and (3) are equal, follows from Duhamel's Theorem, *Introduction to the Calculus*, Chap. XII, § 8. For, on setting,

$$\alpha_k = \rho_k x_k \Delta S_k, \qquad \beta_k = \bar\rho_k x_k \Delta S_k,$$

we have $$\frac{\beta_k}{\alpha_k} = \frac{\bar\rho_k}{\rho_k}, \qquad \text{and} \qquad \lim_{n=\infty}\frac{\bar\rho_k}{\rho_k} = 1.$$

Thus we obtain the general formula

(4) $$\bar x = \frac{\displaystyle\iint_S \rho\, x\, dS}{M}.$$

Example. To find the centre of gravity of a triangle.

Let the axes be chosen as in the figure, and let the equations of the sides through the origin be:

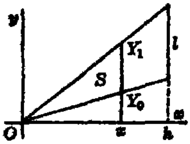


FIG. 11

$$y = \lambda_0 x, \qquad y = \lambda_1 x, \qquad \lambda_0 < \lambda_1.$$

Then we have:

$$\iint_S x\, dS = \int_0^h dx \int_{Y_0}^{Y_1} x\, dy,$$

where $$Y_0 = \lambda_0 x, \qquad Y_1 = \lambda_1 x.$$

Hence $$\int_{Y_0}^{Y_1} x\, dy = xy\Big|_{Y_0}^{Y_1} = (\lambda_1 - \lambda_0)x^2,$$

$$\int_0^h dx \int_{Y_0}^{Y_1} x\, dy = \int_0^h (\lambda_1 - \lambda_0)x^2\, dx = (\lambda_1 - \lambda_0)\frac{h^3}{3}.$$

Since $\lambda_1 h - \lambda_0 h = l$, we have finally,

$$\iint_S x\,dS = \tfrac{1}{6}\,lh^2;$$

and since $A = \frac{1}{2}\,lh$, we see that

$$\bar{x} = \tfrac{2}{3}\,h.$$

But O was any vertex, and hence we have the result that the centre of gravity of a triangle lies on a line which is twice as far from a vertex as from the opposite side. The centre of gravity must, therefore, be at the intersection of the medians.

The advantage of the solution by double integrals is its directness. It was not necessary first to develop a special formula to fit this case, as was done in the *Introduction*, Chap. XII, § 10, Exs. 6, 7.

EXERCISES

Find the centre of gravity of each of the following laminas:

1. The lamina of the example studied in the text of § 6.

Ans. $\bar{x} = \frac{4}{5}a$.

2. The lamina of Exercise 1, § 6.

3. A lamina in the form of a 45° right triangle, if the density at any point is proportional to the product of the distances from the two legs.

4. The lamina of Exercise 6, § 6.

5. The lamina of Exercise 7, § 6.

6. Use the present method to obtain the earlier formula,

$$\bar{x} = \frac{\displaystyle\int_a^b xy\,dx}{A},$$

for a region bounded by the axis of x, two ordinates, and the curve $y = f(x) > 0$.

7. Show that, for the plane area of Question 6,

$$\bar{y} = \frac{\displaystyle\int_a^b y^2\,dx}{2A}.$$

8. Moments and Products of Inertia. *By reasoning similar to that of § 7 the following formulas are obtained. The moment of* inertia of a lamina of variable density about the origin is given by the equation:

(1) $$I = \iint_S \rho (x^2 + y^2)\,dS.$$

The moment of inertia of such a lamina about the axis of y is given by the equation:

(2) $$I = \iint_S \rho x^2\,dS;$$

with a similar formula for the moment of inertia about the axis of x.

Products of Inertia. Let $m_1, m_2, \cdots, m_n$ be a system of particles, whose coordinates are $(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)$, respectively. By the *product of inertia* of this system with respect to the coordinate axes is meant the sum:

(3) $$P = \sum_{i=1}^{n} m_i x_i y_i = m_1 x_1 y_1 + m_2 x_2 y_2 + \cdots + m_n x_n y_n.$$

The definition is extended to continuous distributions by the usual methods of the calculus. Thus for a lamina

(4) $$P = \iint_S \rho xy\,dS.$$

By the *product of inertia of a plane area* is meant the value of the above integral when $\rho = 1$. The product of inertia with respect to two parallels to the coordinate axes, which intersect in the point (a, b), is defined as

(5) $$\iint_S \rho (x - a)(y - b)\,dS.$$

EXERCISES

Determine the following moments of inertia by double integration.

1. A square about its centre. *Ans.* $\frac{1}{6} Ma^2$.

2. A square about a side. *Ans.* $\frac{1}{3} Ma^2$.

3. A right triangle about a vertex.

4. A right triangle about the right angle.

5. A segment of a parabola cut off by the latus rectum, about the focus.

6. The same, about the latus rectum.

7. A uniform lamina bounded by the parabola $y^2 = 4\,ax$, the line $x + y = 3\,a$, and the axis of x, about the axis of y. Work the problem both ways, integrating first with regard to x, then with regard to y; and then in the opposite order. *Ans.* $I = \frac{48}{7}\rho a^4$.

8. Give the details of the proof by which formula (1) is established.

Determine the following products of inertia.

9. A square, two of whose sides lie along the positive coordinate axes.

10. Each of the plane areas shown herewith.

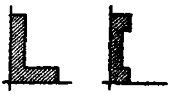


Fig. 12

11. Show that, if either the axis of x or that of y is an axis of symmetry, the product of inertia with respect to the coordinate axes vanishes.

9. Theorems of Pappus. Theorem I. *If a closed plane curve rotate about an external axis lying in its plane, the volume of the ring thus generated is the same as that of a cylinder whose base is the region S inclosed by the curve and whose altitude is the distance through which the centre of gravity of S has traveled:*

$$(1) \qquad V = 2\,\pi h \cdot A,$$

where h denotes the distance of the centre of gravity of S from the axis, and A, the area of S.

Let the axis of rotation be taken as the axis of y, and let the region S lie in the first quadrant. Divide S up into n elementary regions, and consider the volume, ΔV_k, of the slender ring generated by the k-th of these regions. Think of this ring as made of gutta percha; cut it through along a meridian plane, and straighten it out. Obviously, the part of the gutta percha that was near the axis will be stretched, while the part remote from the axis will be compressed. Hence a cylinder of the same cross-section as the ring and of the same volume, ΔV_k, will have an altitude intermediate between the shortest and the longest parallel of latitude of the ring, or

$$2\,\pi x'_k \Delta S_k < \Delta V_k < 2\,\pi x''_k \Delta S_k,$$

where x'_k, x''_k denote respectively the shortest and the longest distance of any point in ΔS_k from the axis of y.

It is, therefore, possible to choose an x_k intermediate between x'_k and x''_k, and such that

$$\Delta V_k = 2\,\pi x_k \Delta S_k.$$

Hence $$V = \sum_{k=1}^{n} 2\pi x_k \Delta S_k = \lim_{n=\infty} \sum_{k=1}^{n} 2\pi x_k \Delta S_k,$$

or $$V = 2\pi \iint_S x\,dS.$$

On the other hand, by § 7, (1)

$$\iint_S x\,dS = \bar{x}A,$$

and hence $$V = 2\pi hA,$$ Q. E. D.

If the region S rotate only through an angle Θ, the volume V then generated is obviously

(2) $$V = \Theta hA.$$

THEOREM II. *If a plane curve, closed or not closed, rotate about an axis not cutting it and lying in its plane, the area of the surface thus generated is the same as that part of the cylindrical surface having the given curve as generatrix and its elements perpendicular to the plane of the curve, which lies between two parallel planes which are perpendicular to these elements and whose distance apart is the distance traversed by the centre of gravity of the curve:*

$$S = 2\pi h \cdot l \qquad or \qquad \Theta h \cdot l.$$

The proof is similar to that of the first theorem.

EXERCISES

1. Find the volume of a torus, or anchor ring.

2. Obtain the centre of gravity of a semi-circular lamina, assuming the formula for the volume of a sphere.

3. Obtain the centre of gravity of a semi-circular wire.

4. Find the area of an anchor ring.

5. Prove Theorem II.

10. The Iterated Integral in Polar Coordinates. We have computed the volume V under the surface $z = f(x, y)$ by iterated integration, using Cartesian coordinates. Let us now compute the same volume, using polar coordinates. To do this, we divide the solid up into thin wedge-shaped slabs (the slab not extending in general clear to the edge of the wedge) by means of n equally spaced planes through the axis of z: $\theta = \theta_0 = \alpha,\ \theta_1, \cdots,\ \theta_n = \beta$, and approximate *to the volume of the k-th slab, ΔV_k, as follows.* Let A_k be the area of the section of the plane $\theta = \theta_{k-1}$ with the solid, and let this section rotate about the axis of z through the angle $\Delta\theta$. Then, by § 9, Theorem I, the volume generated is $\Delta\theta \cdot h_k A_k$, and the sum of such volumes,

$$\sum_{k=1}^{n} h_k A_k \Delta\theta,$$

is a good approximation for V. In fact, when we visualize the totality of these pieces, we see that the volume of the solid thus obtained approaches V as its limit, when $n = \infty$. Hence

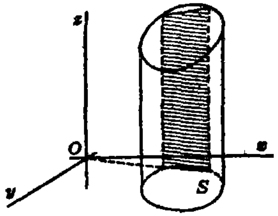


FIG. 13

(1) $$V = \lim_{n=\infty} \sum_{k=1}^{n} h_k A_k \,\Delta\theta = \int_{\alpha}^{\beta} hA\,d\theta.$$

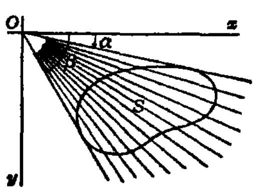


FIG. 14

The product hA which forms the integrand corresponds to the cross-section made by an arbitrary plane $\theta = \theta'$. Writing the equation of the surface in the form

$$z = F(r, \theta)$$

and recalling the general formula for the abscissa of the centre of gravity G of a plane area, § 7, Ex. 6, we see that here the coordinates in the meridian plane $\theta = \theta'$ are r and z, corresponding respectively to the x and y of the above formula. Moreover, $h = \bar{x}$, $R' = a$, and $R'' = b$, where R' and R'' are obtained by cutting the region S by the ray $\theta = \theta'$ and taking, on the line-segment intercepted by S, the smallest r for R' and the largest r for R''. Thus we obtain:

$$hA = \int_{R'}^{R''} rF(r, \theta')\,dr.$$

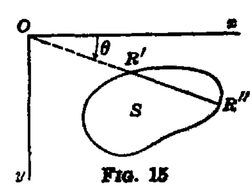


FIG. 15

Substituting this last expression in (1) and dropping the accents, we get the final formula:

$$V = \int_{\alpha}^{\beta} d\theta \int_{R'}^{R''} rF(r, \theta)\,dr,$$

Since the value of the double integral is also equal to V, the Fundamental Theorem of § 4 now takes on the following form.

THEOREM:

$$(2) \qquad \iint\limits_S F(r, \theta)\, dS = \int\limits_{\alpha}^{\beta} d\theta \int\limits_{R'}^{R''} rF(r, \theta)\, dr,$$

where the limits of integration, namely, R' and R'' on the one hand, and α and β on the other, are determined as set forth above.

It is particularly to be borne in mind that, before the first integration is performed, the integrand of the double integral, $F(r, \theta)$, must be multiplied by r to give the integrand, $rF(r, \theta)$, of the iterated integral.

The Inverse Order of Integration. If instead of using the planes $\theta = \theta_0, \theta_1, \cdots, \theta_n$ we had divided the solid up by the cylinders $r = r_0 = a, r_1, \cdots, r_n = b$, we should have been led to the result:

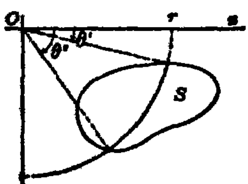


FIG. 16

$$(3) \qquad \iint\limits_S F(r, \theta)\, dS = \int\limits_a^b dr \int\limits_{\Theta'}^{\Theta''} rF(r, \theta)\, d\theta.$$

Here, the first integration is performed on the supposition that r is held fast and that θ varies algebraically from the smallest value, Θ', which it has in S corresponding to the given value of r to the largest value, Θ''.

Example 1. To find the moment of inertia of a uniform circular disc about a diameter.

Let the equation of the bounding circle be

$$x^2 + y^2 = a^2, \qquad \text{or} \qquad r = a,$$

and let the diameter lie in the axis of y. Then

$$I = \rho \iint\limits_S x^2\, dS,$$

extended over the circle.

The integrand, x^2, of the double integral becomes in polar coordinates

$$F(r, \theta) = r^2 \cos^2\theta.$$

The integrand of the iterated integral is, therefore,

$$rF(r, \theta) = r^3 \cos^2\theta.$$

Thus we have

$$I = \rho \int_0^{2\pi} d\theta \int_0^a r^3 \cos^2\theta \, dr.$$

The first integration gives

$$\frac{r^4}{4}\cos^2\theta \Big|_0^a = \frac{a^4}{4}\cos^2\theta.$$

Hence
$$I = \frac{\rho a^4}{4}\int_0^{2\pi} \cos^2\theta \, d\theta = \frac{\rho a^4}{8}(\theta + \sin\theta\cos\theta)_0^{2\pi} = \frac{\pi\rho a^4}{4} = \frac{Ma^2}{4}.$$

Example 2. The density of a square lamina is proportional to the distance from one corner. Find its mass.

Here,
$$M = \int_S\int \rho \, dS, \qquad \rho = cr.$$

Clearly, it is sufficient to compute the mass of one of the right triangles into which the square can be divided. Thus

$$\tfrac{1}{2}M = c\int_0^{\frac{\pi}{4}} d\theta \int_0^R r^2 \, dr, \qquad R = \frac{a}{\cos\theta}.$$

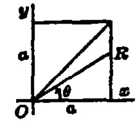


FIG. 17

$$\int_0^R r^2\, dr = \frac{a^3}{3\cos^3\theta},$$

$$\int_0^{\frac{\pi}{4}} \frac{d\theta}{\cos^3\theta} = \frac{\sin\theta}{2\cos^2\theta} + \tfrac{1}{2}\log\tan\left(\frac{\pi}{4}+\frac{\theta}{2}\right)\Big|_0^{\frac{\pi}{4}} = \frac{1}{\sqrt{2}} + \tfrac{1}{2}\log\tan\frac{3\pi}{8}.$$

Hence, finally:

$$M = \frac{2ca^3}{3}\left(\frac{1}{\sqrt{2}} + \tfrac{1}{2}\log\tan\frac{3\pi}{8}\right).$$

EXERCISES

1. Compute the moment of inertia of a uniform circular disc about its centre. *Ans.* $\frac{1}{2}Ma^2$.

2. The density of a circular disc is proportional to the distance from the centre. Find the moment of inertia with respect to the centre, and determine the radius of gyration.
Ans. $I = \frac{2}{5}\pi c a^5$; $k = a\sqrt{\frac{3}{5}}$.

3. Compute the moment of inertia of the square lamina of Example 2 about the point O, and find its radius of gyration.

4. *Find the centre of gravity of the square lamina of Example 2.*

5. Determine the moment of inertia about the origin of the part of the first quadrant which is cut off by two successive coils of the spiral

$$r = e^{\theta},$$

the inner boundary going through the point $\theta = 0$, $r = 1$.

6. How far from the pole, $r = 0$, is the centre of gravity of a lobe of one of the roses,

$$r = a \cos 3\theta \qquad \text{or} \qquad r = a \sin 3\theta?$$

7. Find the moment of inertia of the lemniscate

$$r^2 = a^2 \cos 2\theta$$

about the point $r = 0$.

8. Determine the centre of gravity of one lobe of the lemniscate of Question 7.

9. Give the details of the proof of Formula (3) in the text.

10. Show that the area of any plane region S is given by the formula

$$A = \iint_S dS.$$

Hence, show that the area bounded by the curve

$$\theta = \sin r$$

and the portion of the ray $\theta = 0$ between the pole and the point $r = \pi$ is π. Draw roughly the boundary of the region in question.

11. **Areas of Surfaces.** We have determined the area under a plane curve and the area of a surface of revolution by means of simple integrals. The general problem of finding the area of any curved surface is solved by double integration.

Let the equation of the surface be

$$(1) \qquad z = f(x, y)$$

and let the projection on the (x, y)-plane of the part $\mathfrak{S}$ of this surface whose area A is to be computed, be the region S. Divide S up into elementary areas and erect on the perimeter of each as directrix a cylindrical surface. By means of these cylinders the surface $\mathfrak{S}$ is divided into elementary pieces, of area ΔA_k, $(k = 1, \cdots, n)$, and we next consider how we may approximate to these partial areas. Evidently this may be done by constructing the tangent plane at a point (x_k, y_k, z_k) of the k-th elementary area and computing the area cut out of this plane by the cylinder in question. Now the orthogonal cross-section of this cylinder is of area ΔS_k, and hence the oblique section will have the area

$$\Delta S_k \sec \gamma_k,$$

where γ_k is the angle between the planes, or between their normals. The desired approximation is thus seen to be

$$\sum_{k=1}^{n} \Delta S_k \sec \gamma_k,$$

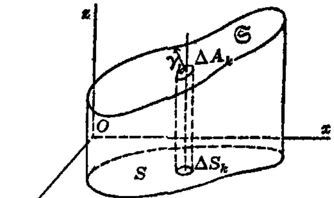


FIG. 18

and consequently A is equal to the limit of this sum, or*

$$(2) \qquad A = \iint_S \sec \gamma \, dS,$$

where γ denotes the acute angle between the normal to the surface and the axis of z.

There are three leading forms for representing the surface analytically (cf. Chap. VI, § 1), namely, the explicit form (1), the implicit form

$$(3) \qquad F(x, y, z) = 0,$$

and the parametric form

$$(4) \qquad x = f(u, v), \qquad y = \phi(u, v), \qquad z = \psi(u, v),$$

where f, ϕ, ψ have continuous partial derivatives of the first order and at least one of the two-rowed determinants

$$j_1 = \begin{vmatrix} \frac{\partial \phi}{\partial u} & \frac{\partial \phi}{\partial v} \\ \frac{\partial \psi}{\partial u} & \frac{\partial \psi}{\partial v} \end{vmatrix}, \qquad j_2 = \begin{vmatrix} \frac{\partial \psi}{\partial u} & \frac{\partial \psi}{\partial v} \\ \frac{\partial f}{\partial u} & \frac{\partial f}{\partial v} \end{vmatrix}, \qquad j_3 = \begin{vmatrix} \frac{\partial f}{\partial u} & \frac{\partial f}{\partial v} \\ \frac{\partial \phi}{\partial u} & \frac{\partial \phi}{\partial v} \end{vmatrix},$$

— here, j_3, — is different from 0.

Corresponding to the form (1) formula (2) becomes

$$(5) \qquad A = \iint_S \sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2} \, dS.$$

* It is a fundamental principle of elementary geometry to refer all geometrical truth back directly to the definitions and axioms. What are the axioms on which this formula depends? The answer is: *The formula itself is an axiom.* The justification for this axiom is the same as for any other physical law, namely, that the physical science, here geometry, built on it is in accord with experience.

The second form gives

(6)
$$A = \iint_S \frac{\sqrt{\left(\frac{\partial F}{\partial x}\right)^2 + \left(\frac{\partial F}{\partial y}\right)^2 + \left(\frac{\partial F}{\partial z}\right)^2}}{\left|\frac{\partial F}{\partial z}\right|}\, dS.$$

The parametric form, (4), yields a formula, the complete establishment of which must be deferred till the transformation of double integrals has been taken up in Chap. XII, § 4. It is:

(7)
$$A = \iint_\Sigma \Delta\, du\, dv,$$

where
$$\Delta = \sqrt{j_1^2 + j_2^2 + j_3^2}$$

and Σ is the part of the (u, v)-plane which corresponds to the surface $\mathfrak{S}$.

Example. Two equal cylinders of revolution are tangent to each other externally along a diameter of a sphere, whose radius is double that of the cylinders. Find the area of the surface of the sphere interior to the cylinders.

It is sufficient to compute the area in the first octant and multiply the result by 8. We have to extend the integral (2) over the region S indicated in Fig. 20. Here,

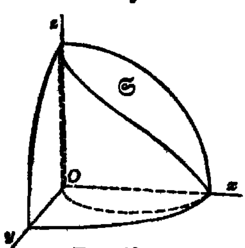


FIG. 19

$$x^2 + y^2 + z^2 = a^2,$$

and by (6)

$$\sec\gamma = \frac{a}{z} = \frac{a}{\sqrt{a^2 - r^2}}, \qquad r^2 = x^2 + y^2.$$

Since the integrand, $\sec\gamma$, depends in a simple way on r, it will probably be well to use polar coordinates in the iterated integral. We have, then:

$$\tfrac{1}{8} A = \iint_S \sec\gamma\, dS = \int_0^{\frac{\pi}{2}} d\theta \int_0^{a\cos\theta} \frac{ar\,dr}{\sqrt{a^2 - r^2}},$$

$$\int_0^{a\cos\theta} \frac{ar\,dr}{\sqrt{a^2 - r^2}} = -a\sqrt{a^2 - r^2}\,\Big|_0^{a\cos\theta} = a^2(1 - \sin\theta),$$

$$\therefore \tfrac{1}{8} A = a^2 \int_0^{\frac{\pi}{2}} (1 - \sin\theta)\, d\theta = a^2\left(\frac{\pi}{2} - 1\right),$$

$$A = 4\pi a^2 - 8a^2.$$

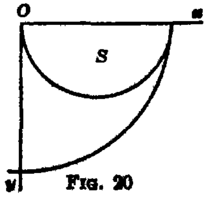


FIG. 20

Objection may be raised to the foregoing solution on the ground that the integrand, $\sec\gamma = a/\sqrt{a^2 - r^2}$, does not remain finite throughout S, but becomes infinite at the point $\theta = 0$, $r = a$. We may avoid this difficulty by computing first only so much of the area as lies over the angle $\alpha \leq \theta \leq \pi/2$, where the positive quantity α is chosen arbitrarily small. The value of this area is

$$a^2 \int_{\alpha}^{\frac{\pi}{2}} (1 - \sin\theta)\, d\theta = a^2\left(\frac{\pi}{2} - \alpha - \cos\alpha\right),$$

and its limit, when α approaches 0, is $a^2(\frac{1}{2}\pi - 1)$.

EXERCISES

1. A cylinder is constructed on a single loop of the curve

$$r = a\cos n\theta$$

as generatrix, its elements being perpendicular to the plane of this curve. Determine the area of the portion of the sphere

$$x^2 + y^2 + z^2 = 2\,az$$

which the cylinder intercepts. *Ans.* $\dfrac{2(\pi - 2)a^2}{n}$

2. Compute the area of the surface

$$z = x + y^2$$

which lies above the triangle of the (x, y)-plane whose sides lie along the lines

$$x = 0, \qquad y = x, \qquad y = a.$$

3. A column is bounded by the four vertical planes

$$x = 0, \qquad x = a, \qquad y = 0, \qquad y = a;$$

the horizontal plane, $z = 0$; and the surface

$$z = 1 + 2x + 3y + x^2.$$

Find the total area of its surface.

4. Determine the area of the surface

$$z = xy$$

included within the cylinder

$$x^2 + y^2 = a^2$$

5. A cylindrical surface is erected on the curve $r = \theta$ as generatrix, the elements being perpendicular to the plane of this curve. Find the area of the portion of the surface

$$z = xy$$

which is bounded by the (y, z)-plane and so much of the cylindrical surface as corresponds to $0 \leqq \theta \leqq \pi/2$.

6. The horizontal cylinder

$$z = f(x)$$

is cut by a vertical cylinder whose base is the region S of the (x, y)-plane bounded by the lines $x = a$, $x = b$ and the curves

$$y = \phi(x), \qquad y = \psi(x),$$

where $\phi(x)$ and $\psi(x)$ are two functions continuous in the interval $a \leqq x \leqq b$, and

$$\phi(x) < \psi(x), \qquad a < x < b.$$

Show that the area cut out of the cylinder is given by the formula

$$A = \int_a^b (Y_1 - Y_0)\sqrt{1 + \left(\frac{dz}{dx}\right)^2}\, dx, \tag{8}$$

where $\qquad Y_0 = \phi(x), \qquad Y_1 = \psi(x).$

7. Two cylinders of revolution, of equal radii, intersect, their axes cutting each other at right angles. Show that the total area of the surface of the solid included within these cylinders is $16\,a^2$.

8. Obtain formula (8) directly, without the use of double integrals.

9. Show that the lateral area of that part of either of the cylinders of the Example of the text, which is contained in the sphere, is $4\,a^2$.

10. From formula (7) deduce the formula for the area of a surface of revolution:

$$A = 2\pi \int_a^b y\sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx.$$

Write the equations of the surface in the parametric form:

$$x = u, \qquad y = f(u)\cos v, \qquad z = f(u)\sin v.$$

Thus

$$A = \iint_\Sigma f(u)\sqrt{1 + f'(u)^2}\, du\, dv = 2\pi \int_a^b f(u)\sqrt{1 + f'(u)^2}\, du.$$

11. Express the equations of a torus parametrically, and thus find its area.

12. Fluid Pressures. Let a vertical plane area be immersed in a liquid as described in the *Introduction to the Calculus*, Chap. XII, § 11. The pressure, P, of the liquid on the surface can be formulated as a double integral by means of the same physical considerations as those of the earlier deduction in terms of a simple integral. We find:

$$(1) \qquad P = w\int_S\int (x+c)\,dS.$$

Specific Pressure. We have hitherto dealt only with the total pressure on a surface, and this is a *force* in the ordinary sense of the term — a push or a pull. In hydromechanics one meets the conception of the *pressure at a point*, and by this is meant the following. Consider an arbitrary point, Q, in the fluid; pass a surface through Q, and draw a small closed curve C on the surface, which shall include Q in its interior or on its boundary. Let P denote the pressure of the fluid on the part of the surface enclosed by C, and let A be the area of this small piece.* Then the ratio P/A represents the *average pressure* on the piece, and the limit, p, approached by the average pressure is what is meant by the *pressure at the point*, Q, or the *specific* pressure:

$$p = \lim \frac{P}{A}.$$

The direction of P approaches as its limit the normal at Q.

It is shown in hydromechanics that the limit p is the same for all surfaces through Q, no matter what their form and what their orientation may be. This fact is often stated briefly in the words: "fluids press equally in all directions."

The total pressure on one side of a plane surface, S, immersed in a fluid and oriented in any way, is normal to the surface and is given by the formula:

$$P = \int_S\int p\,dS.$$

The proof is similar to that of equation (5), § 6, and is left to the student.

* There are, of course, two sides of this small piece of surface, and two pressures (equal and opposite), one on each side. We fix our attention on one of these sides, and on the pressure exerted by the fluid on it. A draftsman's thumb-tack is a suggestive model of the surface and the force.

Example. The density of a certain liquid, acted on only by the force of gravity, is proportional to the depth below the surface. Find the pressure at an arbitrary point.

Let x denote the depth below the surface. Then

$$\rho = cx.$$

The weight of a vertical column of the liquid, ξ units high and of cross-section of area A, is seen to be

$$A\int_0^{\xi} \rho\, dx = \tfrac{1}{2} cA\xi^2,$$

forces being measured in gravitational units.

If Q be an arbitrary point of the liquid, at depth x, and if the surface through Q be taken as a horizontal plane, then the average pressure on this surface will be

$$\frac{P}{A} = \frac{\tfrac{1}{2} cAx^2}{A} = \tfrac{1}{2} cx^2.$$

Hence the pressure at Q (*i.e.* the *specific* pressure) will be: $p = \tfrac{1}{2} cx^2$.

EXERCISES

1. Give the proof of the theorem embodied in formula (1), drawing the requisite figure.

2. By means of the above theorem show that the pressure P is equal to the weight of a cylindrical column of the liquid, whose cross-section is the area S, and whose altitude is the distance of the centre of gravity of S below the surface of the liquid.

3. From (1) deduce the formula for P given in the *Introduction*, Chap. XII, § 11, p. 315, (4).

4. The density of a certain liquid, acted on by gravity, is proportional to the distance below the surface. Prove that the pressure on a vertical rectangle, one side of which is in the surface, is equal to the weight of a vertical column of the liquid standing on an equal rectangle and extending to a depth of $\sqrt{\tfrac{1}{3}}$ times the length of the vertical sides of the original rectangle.

5. Find the centre of pressure in the preceding problem.

6. Use the results of the present paragraph to obtain a simpler solution of Ex. 7, p. 316, of the *Introduction*.

13. Attractions. Let it be required to find the attraction of a material plane surface, or lamina, of variable density on a particle of unit mass situated in its plane, but external to the area. The problem will be solved if we can find the component of the resultant attraction along an arbitrary direction. For then we can compute the components along two different directions, and by the law of the parallelogram of forces find the resultant attraction.

Divide the area up into sub-regions and approximate to the component attraction of each of these. It will be convenient to choose our coordinates with the particle at the origin, and to compute the attraction along the axis of x. Since the choice of the direction of the axis of x is arbitrary, our solution is general.

We may take it as physically evident that, when the area S is divided into a large number of small areas, ΔS_k, and the mass, $\Delta M_k = \bar{\rho}_k \Delta S_k$, contained in each is concentrated at one of its points, the attraction of this system of n particles is approximately equal to the attraction of the actual lamina, and that the approximation grows better and better as the number increases, the maximum diameter of the sub-regions approaching 0, so that the limit approached by the attraction of the particles is precisely the attraction of the lamina.

Let (r_k, θ_k) be the polar coordinates of the k-th particle. For convenience we will choose this point so that the value of ρ in it will be equal to $\bar{\rho}_k$; in other words, we will choose it as a point of the locus $\rho = \bar{\rho}_k$. We may now write simply ρ_k instead of $\bar{\rho}_k$, since we are concerned merely with the value of ρ in the point (r_k, θ_k).

The component attraction of the k-th particle along the axis of x will then be

$$K \frac{\rho_k \Delta S_k}{r_k^2} \cos \theta_k.$$

The limit of the sum of these components is the component attraction, F, of the lamina along that direction, or

(1) $$F = \lim_{n=\infty} \sum_k K \frac{\rho_k \Delta S_k}{r_k^2} \cos \theta_k.$$

But this last limit is a double integral, and hence

(2) $$F = K \int_S\int \frac{\rho \cos \theta \, dS}{r^2}.$$

The variables r and θ have been conveniently described as "polar coordinates," but this does not preclude the use of other systems of coordinates in evaluating the integral. Essentially, the variable r means the distance of a point P of S from the particle at O, and θ means the angle from the direction along which forces are resolved to the line OP.

Example. Find the attraction of a uniform semicircular ring on a particle situated at its centre.

From considerations of symmetry it is clear that the resultant attraction is along the radius perpendicular to the bounding diameter. Hence

$$\tfrac{1}{2}F = K\rho \int_0^{\frac{\pi}{2}} d\theta \int_a^b \frac{\cos\theta}{r}\,dr = K\rho \log\frac{b}{a}.$$

More Refined Treatment. The foregoing physical hypotheses are cruder than is necessary and may be replaced by more refined ones as follows. Since the form of the regions ΔS_k is immaterial, we may take them like the meshes of a spider's web, dividing the region up by circles whose common centre is at O, and by rays, equally spaced, emanating from O. The component, ΔF_k, will then * satisfy the relation

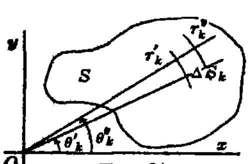


FIG. 21

$$(3)\quad K\frac{\bar\rho_k \Delta S_k}{r_k''^2}\cos\theta_k'' < \Delta F_k < K\frac{\bar{\bar\rho}_k \Delta S_k}{r_k'^2}\cos\theta_k',$$

where r'_k, θ'_k denote the least values of these coordinates in ΔS_k, and r''_k, θ''_k, their greatest.

We can now apply Duhamel's Theorem, *Introduction to the Calculus*, Chap. XII, § 8, setting

$$\alpha_k = K\frac{\rho_k \Delta S_k}{r_k^2}\cos\theta_k, \qquad \beta_k = \Delta F_k,$$

where (r_k, θ_k) is an arbitrary point of ΔS_k, and ρ_k is the value of ρ in this point. Thus formula (2) is established.

It has tacitly been assumed that S lies to the right of the y-axis. If it lies to the left, ΔF_k is negative, and relation (3) is false. However, by making a suitable choice of the points (r'_k, θ'_k) and (r''_k, θ''_k), relation (3) can be reinstated, and the proof proceeds as before; cf. the *Exercises* under Duhamel's Theorem. In case S is cut by the axis of y, each part can be treated as above, the final result being, as before, Formula (2).

* The basis of this statement is our physical intuition. In other words, this statement is precisely the *physical law* which we here postulate. We shall go into this question in more detail presently.

The advantage of this treatment over the earlier one is that the present physical hypotheses are more plausible, — less drastic. They amount substantially to two: First, if a plane distribution of matter (to the right of the y-axis) is replaced by one in which some of the matter of the given distribution has been shoved radially (*i.e.* along rays from O) further away from O, the component of the attraction along the axis of x will thereby be diminished. Secondly, if the given distribution be replaced by one in which some of the matter has been shoved along arcs of circles with the common centre O, so as to recede from the axis of x, the component of the attraction along the axis of x will thereby be diminished.

EXERCISES

1. Determine the attraction of a uniform semicircular lamina on a particle situated in the axis of symmetry and on the circular boundary produced.

2. Determine the attraction of a uniform rectangle on an exterior particle situated in a parallel to two of its sides, passing through its centre.

$$\textit{Ans.}\quad K\frac{mM}{2\,ab}\log\left[\frac{h+a}{h-a}\cdot\frac{b+\sqrt{(h-a)^2+b^2}}{b+\sqrt{(h+a)^2+b^2}}\right].$$

14. Note on Density. Pressure at a Point. Specific Force. We have defined the density of a lamina at an arbitrary point, P, as

$$\rho=\lim\frac{M}{A},$$

where M denotes the mass and A, the area of an arbitrary piece of the lamina containing P, when this piece shrinks down toward P in any manner whatever, its most remote point approaching P as its limit. Thus ρ is seen to be a function of the coordinates (x, y) of P, and we have furthermore required that it be a *continuous* function. This latter hypothesis is, however, superfluous, since the continuity of ρ follows from its definition by means of the above limit.

For, suppose that, at a certain point P_0, ρ were not continuous. Then it would be possible to find a set of points, P_1, P_2, $\cdots$ with P_0 as their limiting point such that the corresponding values ρ_1, ρ_2, $\cdots$ of ρ do not converge toward ρ_0 as their limit. In particular, these points could be so chosen that either

(*i*) $$\rho_k > \rho_0 + h, \qquad k = 1, 2, \cdots,$$
or
(*ii*) $$\rho_k < \rho_0 - h, \qquad k = 1, 2, \cdots,$$

where h denotes a suitably chosen positive constant.

Suppose we have Case (*i*). It is possible to find a region A_k for which

$$\frac{M_k}{A_k}$$

comes as near to ρ_k as one pleases and which, moreover, lies in an arbitrarily small neighborhood of P_k. Moreover, by running a slender spur out from this region so as to include the point P_0 in the extended region, the modified value M_k'/A_k' of the ratio will be at worst but slightly diminished, since A_k' need exceed A_k by only a slight percentage of A_k, and M_k' exceeds M_k. If, then, we choose the first region so that

$$\frac{M_k}{A_k} > \rho_0 + h,$$

the second can also be taken so that

$$\frac{M_k'}{A_k'} > \rho_0 + h.$$

We have here, however, a contradiction. For, the modified regions pertain to the point P_0, and hence, by hypothesis, $\lim M_k'/A_k' = \rho_0$. Thus the theorem we set out to establish is proved, namely, that the continuity of ρ is a consequence of the existence of the limit by means of which ρ was defined.

The theorem and proof hold for three-dimensional distributions of matter, and also for two-dimensional distributions on a curved surface. But it is not true for a one-dimensional distribution. The existence of the density of a material distribution on a wire does not ensure its continuity.

Mean Density. We have assumed that, for any lamina, the equation holds:

(1) $$M = \bar{\rho} A,$$

where $\bar{\rho}$, the *mean density*, is a value of ρ intermediate between the largest value, ρ'', and the smallest value, ρ', which the function ρ takes on in the region:

$$\rho' \leqq \bar{\rho} \leqq \rho''.$$

That this is in fact the case can be shown as follows. Suppose this were not so, and suppose that $\bar{\rho} > \rho''$. Cut the region into two pieces. For these,

$$M_{I} = \bar{\rho}_{I} A_{I}, \qquad M_{II} = \bar{\rho}_{II} A_{II}.$$

At least one of the quantities $\bar{\rho}_I$ and $\bar{\rho}_{II}$, must be as great as $\bar{\rho}$. For otherwise

$$M_I < \bar{\rho} A_I,$$

$$M_{II} < \bar{\rho} A_{II};$$

hence

$$M_I + M_{II} < \bar{\rho}(A_I + A_{II})$$

or

$$M < \bar{\rho} A.$$

But $M = \bar{\rho}A$, and here is a contradiction.

Let A_1 denote one of the regions A_I, A_{II}, for which the average density, $\bar{\rho}_1$, exceeds or at least equals $\bar{\rho}$:

$$\bar{\rho}_1 \geqq \bar{\rho} \qquad \text{in } A_1.$$

Cut A_1 into two pieces and repeat the reasoning. Then, for at least one of these pieces, A_2, we shall have

$$\bar{\rho}_2 \geqq \bar{\rho}_1$$

and hence $\bar{\rho}_2 \geqq \bar{\rho}$. Proceeding in this manner we obtain an infinite sequence of regions, A_1, A_2, $\cdots$, each lying in its predecessor, for each of which

$$\bar{\rho}_k \geqq \bar{\rho}.$$

These regions can be so chosen that their maximum diameter approaches 0 as its limit, and they thus determine a point, P, common to all of them.

The density at P is

$$\rho_P = \lim_{k=\infty} \frac{M_k}{A_k},$$

and since $M_k/A_k = \bar{\rho}_k \geqq \bar{\rho}$, it follows that

$$\rho_P \geqq \bar{\rho} > \rho''.$$

But this is impossible, since ρ'' is the largest value of ρ in the original region.

Pressure at a Point. Body Forces. Similar theorems obtain relating to specific pressure. If a surface, S, form part of the boundary of a fluid, and if Q be a point of S, we have defined the ***pressure at** Q*, or the ***specific pressure** p*, by the limit:

$$p = \lim \frac{P}{A},$$

where A is a portion of S including Q, and P is the pressure on A; cf. § 12. And we have assumed that p is continuous. Reasoning similar to the above shows that p *must* be continuous. Also, it can be proved that the average pressure $\bar{p}$, in case S is plane, cannot exceed the maximum value p'', or cut under the minimum value p', of p in S.

In elasticity and hydromechanics one has to do with *body forces*. A three-dimensional distribution of matter is situated in a field of force. Let P be any point of the mass, and let τ be a *region* including P in its interior or on its boundary. Let the resultant force exerted by the field on the matter in τ be represented by the vector (F). Then $(F)/V$, where V denotes the volume of τ, approaches a limit, the vector $\mathfrak{F}$, the *specific force* — this is our physical hypothesis —

$$\mathfrak{F} = \lim \frac{(F)}{V},$$

when the most remote point of τ approaches P. And now it follows as above that the vector field is *continuous, i.e.* that the vector $\mathfrak{F}$ is continuous. Moreover, an appraisal of the average value of $\mathfrak{F}$ as defined by the equation

$$(F) \doteq V\bar{\mathfrak{F}}$$

is here possible, in terms of inequalities analogous to the foregoing.

In the case of electric and magnetic fields of force the situation is altogether similar.

15. Change of Order of Integration in an Iterated Integral. Hitherto the double integral has come first, and the iterated integral has played the rôle of an agent whereby the double integral is evaluated. We may, however, start out with an iterated integral, as

(1) $$\int_0^a dx \int_0^x (x^2 + y^2)\, dy,$$

and inquire what this integral becomes if the order of integration be reversed.

The question is readily answered by converting the given integral into a double integral. Clearly, the double integral

(2) $$\iint_S (x^2 + y^2)\, dS,$$

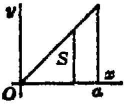


FIG. 22

where S is the region consisting of the triangle indicated in the figure, has the same value as the iterated integral (1). If, now, the double integral be evaluated by means of the iterated integral taken in the other order of integration, we shall have the result we set out to obtain:

$$(3) \qquad \int_0^a dx \int_0^x (x^2 + y^2)\, dy = \int_0^a dy \int_y^a (x^2 + y^2)\, dx.$$

EXERCISES

Express the following iterated integrals as double integrals, and draw a figure showing the region S over which each is to be extended. Reverse the order of integration in each of the iterated integrals.

1. $\displaystyle\int_0^a dx \int_0^b f(x, y)\, dy.$

2. $\displaystyle\int_0^a dx \int_0^{\sqrt{a^2 - x^2}} f(x, y)\, dy.$

3. $\displaystyle\int_0^1 dy \int_{y^2}^1 f(x, y)\, dx.$

4. $\displaystyle\int_{\frac{1}{3}}^{\frac{2}{3}} dy \int_{y^2}^{\sqrt{y}} f(x, y)\, dx.$

5.* $\displaystyle\int_0^{2\pi} d\theta \int_0^a rF(r, \theta)\, dr.$

6. $\displaystyle\int_1^2 dr \int_{-\pi}^{\pi} rF(r, \theta)\, d\theta.$

7. $\displaystyle\int_0^{\frac{1}{4}\pi} d\theta \int_0^{a \sec\theta} rF(r, \theta)\, dr.$

8. $\displaystyle\int_a^{a\sqrt{2}} dr \int_{\cos^{-1}\frac{a}{r}}^{\frac{1}{4}\pi} \Phi(r, \theta)\, d\theta.$

16. Surface Integrals. The extension of the conception of the double integral from a plane region S to a curved surface $\mathfrak{S}$ is immediate. Let a function f be given, defined at each point of $\mathfrak{S}$, and let it be continuous over $\mathfrak{S}$. Let $\mathfrak{S}$ be divided up into a large number of small regions,† $\Delta\mathfrak{S}_k$, and let f_k be the value of f at an arbitrary point of $\Delta\mathfrak{S}_k$. Form the sum:

$$\sum_{k=1}^{n} f_k \Delta\mathfrak{S}_k.$$

* In Exs. 5–8 it is assumed that r, θ are interpreted as polar coordinates. Work these same problems, taking r, θ, as Cartesian coordinates.

† The notation $\Delta\mathfrak{S}_k$, ΔS_k is here used both for the surface and for its area.

17. A New Proof of the Fundamental Theorem. It is possible to deduce the Fundamental Theorem of § 4 without the aid of the geometric concept of the volume V. This method has a two-fold advantage: first, it throws a strong light on the determination of the limits of integration; secondly, it is the only method available when we come to triple integrals.

We shall give only an outline of the method in the present paragraph, our object being to set forth clearly the *thought* and the *technique*. A *proof* will be given in Chap. XII, § 3.

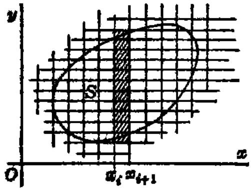


FIG. 23

In the sum:

$$\sum_{k=1}^{n} f(x_k, y_k)\,\Delta S_k, \tag{1}$$

whose limit is the double integral

$$\iint_S f\,dS, \tag{2}$$

we may choose as the sub-regions, or elementary areas, rectangles with sides Δx, Δy, thus making $\Delta S_k = \Delta x\,\Delta y$, and then add all those terms together which correspond to rectangles lying in a column parallel to the axis of y. This partial sum can be represented as follows:

$$\Delta x \sum_{j=1}^{q} f(x_i, y_j)\,\Delta y,$$

where we have assigned new indices, i and j, to the coordinates of the point (x_k, y_k), and where furthermore we have chosen the points (x_k, y_k) of this column so that they all have the same abscissa, x_i.

If, now, holding x_i and Δx fast, we allow q to increase without limit, Δy approaching 0 as its limit, we have

$$\Delta x \lim_{q=\infty} \sum_{j=1}^{q} f(x_i, y_j)\,\Delta y = \Delta x \int_{y_i'}^{y_i''} f(x_i, y)\,dy. \tag{3}$$

Next, we add all the limits of these columns together:

$$\sum_{i=1}^{p} \Delta x \int_{y_i'}^{y_i''} f(x_i, y)\,dy,$$

and allow p to increase without limit, Δx approaching 0 This gives

$$\lim_{p=\infty}\sum_{i=1}^{p}\Delta x\int_{y'_i}^{y''_i} f(x_i, y)\,dy = \int_a^b dx\int_{y'}^{y''} f(x, y)\,dy,$$

i.e. the iterated integral of the Fundamental Theorem.*

This method of deduction is not rigorous, for we have not proven that we get the same result when we take the limit by columns and then take the limit of the sum of the columns, as when we allow all the ΔS_k's to approach 0 simultaneously in the manner prescribed in the definition of the double integral. It is nevertheless useful as giving us additional insight into the structure of the iterated integral, for it enables us to think of the first integration as corresponding to a *summation of the elements in* (1) *by columns*, and of the second integration as corresponding to the *summation of these columns.* Moreover, when we come to polar coordinates in the next paragraph, it helps to explain and make evident the limits of integration, and also the presence of the factor r in the integrand.

18. Continuation; Polar Coordinates. Let the region S be divided up into elementary areas by the circles $r = r_i$, $r_{i+1} - r_i = \Delta r$, and the straight lines $\theta = \theta_j$, $\theta_{j+1} - \theta_j = \Delta\theta$. Then

$$\Delta S_k = r_k \Delta r\,\Delta\theta + \tfrac{1}{2}\,\Delta r^2\,\Delta\theta,$$

and hence, in taking the limit of the sum (1), ΔS_k may, by Duhamel's Theorem, be replaced by $r_k\Delta r\,\Delta\theta$. Writing

$$f(x, y) = F(r, \theta)$$

we have, therefore,

$$\iint_S f\,dS = \lim_{n=\infty}\sum_{k=1}^{n} F(r_k, \theta_k)\,r_k\Delta r\,\Delta\theta.$$

In order to evaluate this latter limit, we may replace (r_k, θ_k) by (r_i, θ_j) and, holding θ_j fast, add together those terms that correspond to elementary areas lying in the angle between the rays $\theta = \theta_j$ and $\theta = \theta_{j+1}$, thus getting

$$\Delta\theta\sum_{i=1}^{p} F(r_i, \theta_j)\,r_i\,\Delta r.$$

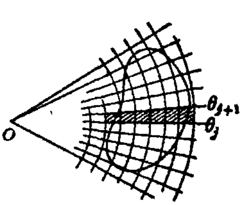


FIG. 24

* The limits of integration, y' and y'', are the functions denoted in § 3 by Y_0 and Y_1. The change in notation was due to the integral in (3), where y'_i would have been awkwardly expressed as $(Y_0)_i$, and likewise for y''_i. A similar change of notation is made below in § 18.

The limit of this sum, as $p = \infty$, is

$$\Delta\theta \int_{r_j'}^{r_j''} F(r, \theta_j)\, r\, dr.$$

Next, add all the limits thus obtained for the successive elementary angles together and take the limit of this sum. We thus get

$$\lim_{q=\infty} \sum_{j=1}^{q} \Delta\theta \int_{r_j'}^{r_j''} F(r, \theta_j)\, r\, dr = \int_{\alpha}^{\beta} d\theta \int_{r'}^{r''} F(r, \theta)\, r\, dr,$$

i.e. the first iterated integral of § 10.

If on the other hand we hold r_i fast and add the terms that correspond to elementary areas lying in the circular ring bounded by the radii $r = r_i$ and $r = r_{i+1}$, we get

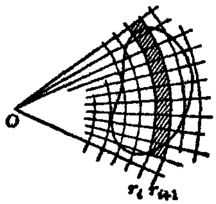


FIG. 25

$$\Delta r \sum_{j=1}^{q} F(r_i, \theta_j)\, r_i\, \Delta\theta,$$

and the limit of this sum, when $q = \infty$, is

$$\Delta r \int_{\theta_i'}^{\theta_i''} F(r_i, \theta)\, r_i\, d\theta = r_i\, \Delta r \int_{\theta_i'}^{\theta_i''} F(r_i, \theta)\, d\theta.$$

Adding all these latter limits together and taking the limit of this sum, we have:

$$\lim_{p=\infty} \sum_{i=1}^{p} r_i\, \Delta r \int_{\theta_i'}^{\theta_i''} F(r_i, \theta)\, d\theta = \int_{a}^{b} r\, dr \int_{\theta'}^{\theta''} F(r, \theta)\, d\theta,$$

i.e. the second iterated integral of § 10.

The student may safely use the method of these paragraphs in practice, since the ideas involved are all correct as far as they go. A fallacy arises, however, when the ideas are set forth as if the double integral, written in the form

$$\int\!\!\int_S f(x, y)\, dx\, dy,$$

were the same thing as the iterated integral*

$$\int_a^b dx \int_{y'}^{y''} f(x, y)\, dy,$$

"since each is a sum of infinitesimals, and the order of the summation is immaterial." It is curious that some men who consider themselves practical and object strongly to anything theoretical in mathematics, find no difficulty in accepting as sound a theory which the race has long since outgrown. All the advantages of those earlier attempts to regard the infinitesimal as the ultimate basis of the calculus — like the atoms or electrons of modern physics — can be preserved by recognizing that we have here only the outline of a method; a very suggestive and altogether correct outline, but one which must be filled in by mathematical proof. Such a proof is given below in Chap. XII.

EXERCISES ON CHAPTER III

1. Find the volume cut out of the first octant by the cylinders

$$z = 1 - x^2, \qquad x = 1 - y^2. \qquad \textit{Ans.}\ \tfrac{18}{35}.$$

2. Compute the value of the integral:

$$\iint_S e^{x^2+y^2}\, dS,$$

extended over the interior of the circle

$$x^2 + y^2 = 1. \qquad \textit{Ans.}\ 5.40.$$

3. Evaluate

$$\iint_S (x^2 - 3\,ay)\, dS,$$

where S is a square with its vertices on the coordinate axes, the length of its diagonal being $2\,a$. *Ans.* $\frac{1}{3}a^4$.

4. Express as an iterated integral in polar coordinates the double integral

$$\iint_S f\, dS,$$

extended over a right triangle having an acute angle in the pole. Give both orders of integration.

5. The curve

$$\cos\theta = 3 - 3r + r^2$$

rotates about the initial line. Find the volume of the solid generated. *Ans.* $\frac{28}{15}\pi$.

6. Find the volume cut from a circular cylinder whose axis is parallel to the axis of z, by the (x, y)-plane and the surface

$$xy = az,$$

if the cylinder does not cut the coordinate axes. *Ans.* $\frac{\pi hkr^2}{a}$.

7. A cone of revolution has its vertex in the surface of a sphere, its axis coinciding with a diameter. Find the volume common to the two surfaces. *Ans.* $\frac{4}{3}\pi a^3(1 - \cos^4\alpha)$.

8. Find the volume of a column capped by the surface

$$z = xy,$$

the base of the column being the portion of the first quadrant in the (x, y)-plane which lies between two successive coils of the logarithmic spiral, beginning with $\theta = 0$:

$$r = ae^{\theta}.$$

Ans. $\frac{1}{80}a^4(e^{8\pi} - 1)(e^{2\pi} + 1)$.

9. Find the abscissa of the centre of gravity of the above column.

10. A square hole $2b$ on a side is bored through a cylinder of radius a, the axis of the hole intersecting the axis of the cylinder at right angles. Find the volume of the chips cut out.

Ans. $4b^2\sqrt{a^2 - b^2} + 4a^2b\sin^{-1}\frac{b}{a}$.

11. A square hole $2b$ on a side is bored through a sphere of radius a, the axis of the hole going through the centre of the sphere. Find the volume of the chips cut out.

Ans. $\frac{8}{3}b^2\sqrt{a^2 - 2b^2} + (8a^2b - \frac{8}{3}b^3)\sin^{-1}\frac{b}{\sqrt{a^2 - b^2}} - \frac{8}{3}a^3\sin^{-1}\frac{b^2}{a^2 - b^2}$.

12. Find the area of that part of the surface

$$z = \tan^{-1}\frac{y}{x}$$

which lies in the first octant below the plane $z = \pi/2$ and within the cylinder $x^2 + y^2 = 1$.

13. The density of a square lamina is proportional to the distance from one corner. Determine the mass of the lamina.

Ans. $.765\,\lambda a^3$.

14. Find the centre of gravity of the lamina in the preceding question.

$$Ans.\quad \bar{x}=\bar{y}=a\frac{7\sqrt{2}-2+3\log(1+\sqrt{2})}{8[\sqrt{2}+\log(1+\sqrt{2})]}.$$

15. Obtain a formula for the centre of gravity of a curved surface of variable density.

16. Find the moment of inertia about the origin of the portion of the first quadrant bounded by the curve

$$(x+1)(y+1)=4,$$

correct to three significant figures.

17. Find the moment of inertia of an anchor ring about its axis.

$$Ans.\quad M(\tfrac{3}{4}a^2+b^2).$$

18. Two circles are tangent to each other internally. Determine the moment of inertia of the region between them, about the point of tangency.

19. Find the attraction of a uniform circular disc on a particle situated in a line perpendicular to the plane of the disc at its centre.

20. Solve the same problem for a rectangular disc.

$$Ans.\quad K\frac{mM}{ab}\tan^{-1}\frac{ab}{h\sqrt{h^2+a^2+b^2}}.$$

21. Show that the force with which a homogeneous piece of the surface of a sphere lying wholly in one hemisphere and symmetrical with reference to the diameter perpendicular to the base of the hemisphere attracts a particle situated at the centre of the sphere is proportional to the projection of the piece on the base.

22. Compute the attraction of a homogeneous hemisphere on a particle situated at the point of the spherical surface most remote from the solid.

23. Show that the residual area of the sphere of Question 11 is

$$16a^2\left[\sin^{-1}\frac{a}{\sqrt{2(a^2-b^2)}}-\frac{b}{a}\tan^{-1}\frac{b}{\sqrt{a^2-2b^2}}\right].$$

24. Express as a double integral the iterated integral

$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}d\theta\int_0^{2a\cos\theta}fr\,dr$$

and state over what region the latter is extended.

25. The same for

(a) $$\int_{\beta}^{\frac{\pi}{2}} d\theta \int_{0}^{b \csc \theta} fr\,dr; \qquad (b) \int_{0}^{2a} dy \int_{\frac{y^2}{2a}}^{\sqrt{2ay}} f\,dx.$$

26. Change the order of integration in the following integrals:

(a) $$\int_{0}^{1} dx \int_{x^2}^{1} f(x, y)\,dy; \qquad (b) \int_{0}^{a} dy \int_{\sqrt{a^2-y^2}}^{y+a} f(x, y)\,dx.$$

27. The intensity of light issuing from a point source is inversely proportional to the square of the distance from the source. Formulate as an integral the total illumination of a plane region by an arc light exterior to the plane.

28. Compute the illumination in the foregoing question on the interior of the curve.

$$r^2 = 1 - \theta^2,$$

the light being situated in the perpendicular to the plane of the curve at $r = 0$. *Ans.* $2\,\lambda(1 - h \cot^{-1} h)$.

29. One loop of the curve

$$r^3 = a^3 \cos 3\,\theta$$

is immersed in a liquid, the pole being at the surface and the initial line vertical and directed downward. Find the pressure on the surface. *Ans.* $\dfrac{wa^3\sqrt{3}}{8}$.

30. One loop of the lemniscate

$$r^2 = a^2 \cos 2\,\theta$$

is immersed as the loop of the curve in the preceding question. Find the centre of pressure.

Ans. Distance below the surface $= a\sqrt{2}\left(\dfrac{2}{3\,\pi} + \dfrac{1}{4}\right)$.

31. Obtain a formula for the centre of pressure of an arbitrary fluid on a plane area.

32. Prove that, if a specific pressure exists at every point of a plane area immersed in a fluid, this pressure is a continuous function.

33. Develop a formula for the kinetic energy of a material surface of constant or variable density, which is rotating about an axis.

34. Compute the kinetic energy of the surface of a torus rotating about its axis, the density being uniform.

CHAPTER IV

TRIPLE INTEGRALS

1. Definition of the Triple Integral. Let a function of three independent variables, $f(x, y, z)$, be given, continuous throughout a region V of three-dimensional space. Let this region be divided in any manner into small pieces, of volume ΔV_k, and let (x_k, y_k, z_k) be an arbitrary point of the k-th piece. Form the product $f(x_k, y_k, z_k)\Delta V_k$ and add all these products together:

$$\sum_{k=1}^{n} f(x_k, y_k, z_k)\,\Delta V_k. \tag{1}$$

When n is made to grow larger and larger without limit, the greatest diameter of any of the sub-regions, or elementary volumes, approaching 0 as its limit, the sum (1) approaches a limit, and this limit is defined as the *triple* or *volume integral* of the function f, extended throughout the region V:

$$\lim_{n=\infty} \sum_{k=1}^{n} f(x_k, y_k, z_k)\,\Delta V_k = \iiint_V f\,dV. \tag{2}$$

It is not essential that the totality of the elementary volumes should just fill out the region V. (We might, for example, divide space up into small rectangular parallelepipeds, the lengths of whose edges are Δx, Δy, Δz, and consider such as are interior to V, or such as have at least one point of V in their interior or on their boundary. It is this particular division of space that gives rise to the notation:

$$\iiint_V f(x, y, z)\,dx\,dy\,dz.$$

But what is meant is the volume integral as defined above.

The proof involved in the foregoing definition, — namely, the proof that the sum (1) actually approaches a limit, — has to be given along different lines for triple integrals, from what was possible in the case of double integrals. There, we were able to represent the sum

$$\sum_{k=1}^{n} f(x_k, y_k)\Delta S_k$$

by a variable volume which obviously approached a fixed volume as its limit. Here, we should need a four-dimensional space in which to represent geometrically the sum (1). It is necessary, therefore, to fall back on an analytical proof. The proofs of this theorem and the Fundamental Theorem will be taken up in Chapter XII. The theorems themselves, however, are easily intelligible from their analogy with the corresponding theorems for double integrals, and it is our purpose here to state them and to explain their uses.

Duhamel's Theorem holds for triple integrals, as well as for simple and double integrals, and by means of it, when needed, the earlier formulas for mass, centre of gravity, etc., are extended to three dimensional distributions.

EXERCISES

1. Show that the mass of a body, of variable (but continuous) density ρ, is given by the triple integral:

$$M = \iiint\limits_V \rho \, dV.$$

2. Show that the abscissa, $\bar{x}$, of the centre of gravity of the body is given by the formula:

$$\bar{x} = \frac{\iiint\limits_V \rho x \, dV}{M}, \qquad \text{or} \qquad \bar{x} = \frac{\iiint\limits_V x \, dV}{V},$$

in case the density is constant.

3. Show that its moment of inertia about an arbitrary axis is

$$I = \iiint\limits_V \rho r^2 \, dV,$$

where r denotes the distance of a variable point of the body from the axis.

4.* The component F of the attraction of the body, on a particle of unit mass situated at a point O outside the body, along an arbitrary direction is given by the formula:

$$F = K\iiint\limits_V \frac{\rho\cos\psi}{\mathfrak{r}^2}\,dV,$$

where $\mathfrak{r}$ denotes the distance from O to a variable point P of the body, and ψ is the angle which OP makes with the given direction.

5. Show that the kinetic energy of a rigid body, rotating with angular velocity ω about a fixed axis, is

$$\tfrac{1}{2}\,I\omega^2,$$

where I denotes the moment of inertia about the axis.

2. Evaluation of a Triple Integral by Means of an Iterated Integral. In order to compute the value of the volume integral defined in § 1 we introduce an iterated integral. The method is that of Chap. III, §§ 17, 18. Let the region V be divided up by planes parallel to the coordinate planes into rectangular parallelepipeds whose edges are of lengths Δx, Δy, Δz, and let us take as our elements of volume these little solids. Then $\Delta V_k = \Delta x\,\Delta y\,\Delta z$, and the sum (1) of § 1 becomes

$$(3)\qquad \sum_{k=1}^{n} f(x_k, y_k, z_k)\,\Delta x\,\Delta y\,\Delta z.$$

We will select from this sum the terms that correspond to elements situated in a column parallel to the axis of z and add them together, see Fig. 26:

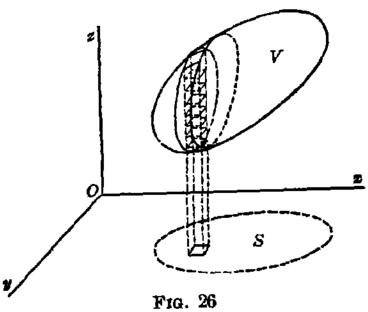


FIG. 26

$$\Delta x\,\Delta y\sum_{l=1}^{s} f(x_i, y_j, z_l)\,\Delta z,$$

where we have assigned new indices, i, j, and l, to the coordinates of the point (x_k, y_k, z_k) and where furthermore we have chosen the

* This exercise may be postponed till the paragraph on the division of space by the surfaces of spherical polar coordinates (§ 3) has been taken up, since this division is the most convenient one for the proof. The object of inserting the exercise at this point is to enable the student who has initiative and imagination to picture to himself this division, after reading the definition of spherical polar coordinates in § 3, and thus to anticipate being shown how to do this thing.

points (x_i, y_j, z_k) of this column so that they all lie in the line $x = x_i$, $y = y_j$. If, now, still holding x_i, y_j, Δx, and Δy fast, we allow s to increase without limit, Δz approaching 0, we have

$$\Delta x \, \Delta y \lim_{s \to \infty} \sum_{l=1}^{s} f(x_i, y_j, z_l) \, \Delta z = \Delta x \, \Delta y \int_{Z_0}^{Z_1} f(x_i, y_j, z) \, dz,$$

where Z_0 is the smallest ordinate of the points of V on the line $x = x_i$, $y = y_j$, and Z_1 is the largest, — we assume for simplicity that the surface of V is met by a parallel to any one of the coordinate axes which traverses the interior of V in two points.

The surface which bounds V consists of two parts, — a lower nappe, represented by the equation $z = \phi_0(x, y)$; and an upper nappe, given by $z = \phi_1(x, y)$. The functions Z_0 and Z_1 have the values respectively:

$$Z_0 = \phi_0(x, y), \qquad Z_1 = \phi_1(x, y).$$

Next, we add all the limits of these columns together:

$$\sum \Phi(x_i, y_j) \, \Delta x \, \Delta y,$$

where we have set

$$\int_{Z_0}^{Z_1} f(x, y, z) \, dz = \Phi(x, y),$$

and take the limit of this sum. The region S of the (x, y)-plane over which this summation is extended consists of the projections of all the points of V on that plane, and hence the limit of this sum is the double integral of $\Phi(x, y)$, extended over S:

$$(4) \qquad \lim \sum \Phi(x_i, y_j) \, \Delta x \, \Delta y = \iint_S \Phi \, dS.$$

We are thus led to the final result:

Fundamental Theorem of the Integral Calculus: *The volume integral* (2) *of* § 1 *is equal to the iterated integral* (4), *or:*

$$(5) \qquad \iiint_V f \, dV = \iint_S dS \int_{Z_0}^{Z_1} f(x, y, z) \, dz$$

The double integral may be evaluated by any of the various iterated integrals studied in Chapter III. If, in particular, the iterated integral in Cartesian coordinates be selected, we have (cf. Fig. 8):

(6) $$\iint_S \Phi(x, y)\, dS = \int_a^b dx \int_{Y_0}^{Y_1} \Phi(x, y)\, dy,$$

where it is assumed that the region S is cut by a parallel to the axis of y at most in two points.

We thus get, as *one* of the final formulas for the volume integral in terms of simple integrals, the following:

(7) $$\iiint_V f\, dV = \int_a^b dx \int_{Y_0}^{Y_1} dy \int_{Z_0}^{Z_1} f(x, y, z)\, dz.$$

Another form in which this iterated integral is written is the following:

$$\int_a^b \int_{Y_0}^{Y_1} \int_{Z_0}^{Z_1} f(x, y, z)\, dz\, dy\, dx.$$

The abbreviated notation

$$\iiint f(x, y, z)\, dx\, dy\, dz$$

may mean either the volume integral or the iterated integral. This notation should be used only when it is explicitly stated, or when it is clear from the context, which is meant.

Example. Find the moment of inertia of a tetrahedron whose face angles at a vertex O are all right angles, about an edge adjacent to O.

Take O as the origin of coordinates and the three adjacent edges as the axes. Then

$$I = \rho \iiint_V (x^2 + y^2)\, dV = \rho \int_0^a dx \int_0^Y dy \int_0^Z (x^2 + y^2)\, dz,$$

where the limits of integration are as follows. First, the limit Z_0 is $= 0$, and the limit $Z_1 = Z$ is the maximum ordinate in V corresponding to an arbitrary pair of values x, y; *i.e.* the ordinate of a point in the oblique face of the tetrahedron:

$$\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1.$$

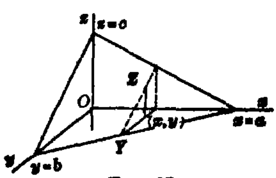


FIG. 27

Hence $$Z = c\left(1 - \frac{x}{a} - \frac{y}{b}\right),$$

and the result of the first integration is:

$$\Phi(x, y) = \int_0^z (x^2 + y^2)\,dz = (x^2 + y^2)\,z \Big|_0^z = c(x^2 + y^2)\left(1 - \frac{x}{a} - \frac{y}{b}\right)$$

$$= c\left[x^2\left(1 - \frac{x}{a}\right) - \frac{x^2}{b}y + \left(1 - \frac{x}{a}\right)y^2 - \frac{y^3}{b}\right].$$

Next, the function $\Phi(x, y)$ must be integrated over the surface S consisting of a triangle bounded by the positive axes of x and y, and the line

$$\frac{x}{a} + \frac{y}{b} = 1.$$

The double integral may be computed by iterated integration, the limits of integration for y being $Y_0 = 0$ and

$$Y_1 = Y = b\left(1 - \frac{x}{a}\right),$$

and those for x being 0 and a. The remainder of the computation is, therefore, as follows:

$$\int_0^Y dy \int_0^z (x^2 + y^2)\,dz = \int_0^Y c\left[x^2\left(1 - \frac{x}{a}\right) - \frac{x^2}{b}y + \left(1 - \frac{x}{a}\right)y^2 - \frac{y^3}{b}\right]dy$$

$$= \frac{bc}{12}\left[6x^2\left(1 - \frac{x}{a}\right)^2 + b^2\left(1 - \frac{x}{a}\right)^4\right],$$

$$\int_0^a dx \int_0^Y dy \int_0^z (x^2 + y^2)\,dz = \int_0^a \frac{bc}{12}\left[6x^2\left(1 - \frac{x}{a}\right)^2 + b^2\left(1 - \frac{x}{a}\right)^4\right]dx$$

$$= \frac{abc}{60}(a^2 + b^2);$$

$$\therefore I = \frac{M(a^2 + b^2)}{10}.$$

The student can verify the answer by slicing the tetrahedron up by planes parallel to the (x, y)-plane and employing the result of Ex. 4 at the end of § 8 in Chap. III, together with the theorem of § 16, Chap. XII, in the *Introduction*.

EXERCISES

1. Find the centre of gravity of the above tetrahedron.

2. Determine the moment of inertia of a rectangular parallelepiped about an axis passing through its centre and parallel to four of its edges.

3. A square column has for its upper base a plane inclined to the horizon at an angle of 45° and cutting off equal intercepts on two opposite edges. How far is the centre of gravity of the column from the axis? *Ans.* $\frac{1}{6}a^2/h$.

4. The density of a cube is proportional to the square of the distance from the centre. Find its mass.

3. Continuation; Spherical Coordinates.* Let P, with the Cartesian coordinates x, y, z, be any point of space. Its spherical coordinates are defined as indicated in the figure. If we think of P as a point of a sphere with its centre at O and of radius r, then θ is the longitude and ϕ is the colatitude of P. We have

$$x = r \sin\phi \cos\theta,$$
$$y = r \sin\phi \sin\theta,$$
$$z = r \cos\phi.$$

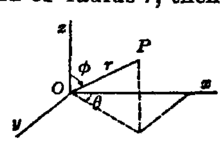


FIG. 28

We propose the problem of computing the volume integral

$$(8) \qquad \lim_{n=\infty} \sum_{k=1}^{n} f(x_k, y_k, z_k)\,\Delta V_k = \iiint_V f\,dV$$

by means of iterated integration in spherical coordinates. For this purpose we will divide the region V up into elementary volumes as follows. Construct (*a*) a set of spheres with O as their common center, $r = r_i$, their radii increasing by Δr; (*b*) a set of half-planes $\theta = \theta_j$, the angle between two successive planes being $\Delta\theta$; and lastly (*c*) a set of cones $\phi = \phi_l$, their semi-vertical angle increasing by $\Delta\phi$: $\phi_{l+1} - \phi_l = \Delta\phi$. The element of volume thus obtained is indicated in Fig. 29.

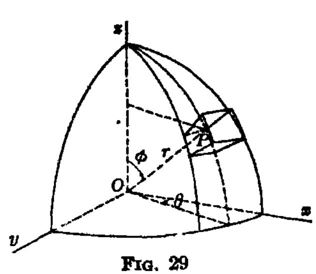


FIG. 29

* Cf. *Analytic Geometry*, Chap XXIV, §§ 1, 2. The student should practice visualizing the loci which are defined by setting one coordinate equal to a constant; then those which arise when two coordinates are held fast.

The lengths of the three edges that meet at right angles at P are Δr, $r\Delta\phi$, $r \sin\phi\ \Delta\theta$, and hence this volume ΔV differs from the volume of a rectangular parallelepiped with the edges just named, or

(9) $$r^2 \sin\phi\, \Delta r\, \Delta\theta\, \Delta\phi$$

by an infinitesimal of higher order:

$$\lim \frac{\Delta V}{r^2 \sin\phi\, \Delta r\, \Delta\theta\, \Delta\phi} = 1.$$

It follows, then, from Duhamel's Theorem that in the limit of the sum (8) we may replace ΔV_k by the infinitesimal (9). If we set

$$f(x, y, z) = F(r, \theta, \phi),$$

we have

$$\iiint\limits_V f\, dV = \lim_{n=\infty} \sum_{k=1}^{n} F(r_k, \theta_k, \phi_k) r_k^2 \sin\phi_k\, \Delta r\, \Delta\theta\, \Delta\phi.$$

Can we evaluate this last limit by iterated integration? It is easy to see that we can. For, the sum is of the type of the sum (3), and hence the method of § 2 is applicable. Following that method, let us select, for example, those terms for which θ and ϕ have a constant value, and add them together:

$$\Delta\theta\, \Delta\phi \sum_{i=1}^{p} F(r_i, \theta_j, \phi_l)\, r_i^2 \sin\phi_l\, \Delta r,$$

where θ_j and ϕ_l are constant. They correspond to elementary volumes lying in a row bounded by the planes $\theta = \theta_j$ and $\theta = \theta_{j+1}$, and by the cones $\phi = \phi_l$ and $\phi = \phi_{l+1}$. Now allow p to increase without limit, Δr approaching 0. This gives, as the limit of the above sum,

$$\Delta\theta\, \Delta\phi \sin\phi_l \int_{R_0}^{R_1} r^2 F(r, \theta_j, \phi_l)\, dr,$$

where R_0 is the distance of the nearest point of V to O on the line $\theta = \theta_j$, $\phi = \phi_l$, and R_1, that of the farthest. We assume for simplicity that the surface of V is met by any one of the lines:

$$\begin{cases} \theta = \text{const.,} \\ \phi = \text{const.,} \end{cases} \qquad \begin{cases} \phi = \text{const.,} \\ r = \text{const.,} \end{cases} \qquad \begin{cases} r = \text{const.,} \\ \theta = \text{const.,} \end{cases}$$

which traverses the interior of V, in two points.

Next, we add all the limits thus obtained together:

(10) $$\sum \Psi(\theta_j, \phi_l)\, \Delta\theta\, \Delta\phi,$$

where we have set

$$\sin\phi \int_{R_0}^{R_1} F(r, \theta, \phi)\, r^2\, dr = \Psi(\theta, \phi),$$

and take the limit of this sum. If we interpret θ and ϕ as the coordinates of a point on the surface of a sphere $r = \text{const.}$ (say, $r = 1$), then the points (θ, ϕ) range over a region S of this spherical surface, which consists of those points in which radii vectores drawn to points of V pierce the surface of the sphere.

The region S is divided into four-sided pieces by the spherical curves $\theta = \theta_i$, $\phi = \phi_l$. The figure strongly suggests the elements which enter into the definition of the double integral. And, in fact, we can identify the limit of the sum (10) with a double integral by transforming the *curved* region S on a *plane* region T as follows. Choose a plane and take a system of Cartesian coordinates (ξ, η) in it. Set

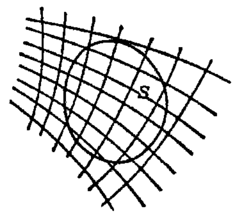


FIG. 30

$$\xi = \theta, \qquad \eta = \phi.$$

Then a point (θ, ϕ) of S goes over into the point (ξ, η) of the plane having the same coordinates, and thus the region S is carried over into a region of the plane. This region we denote by T.

The limit in question now becomes:

$$\text{(11)} \qquad \lim \sum \Psi(\xi_i, \eta_l)\, \Delta\xi\, \Delta\eta = \int\int_T \Psi(\xi, \eta)\, d\xi\, d\eta,$$

i.e. the double integral of the function $\Psi(\xi, \eta)$, extended over the region T. This double integral can be evaluated by means of an iterated integral, as set forth in Chap. III. In particular, we have

$$\text{(12)} \qquad \int\int_T \Psi(\xi, \eta)\, dT = \int_\alpha^\beta d\xi \int_{H_0}^{H_1} \Psi(\xi, \eta)\, d\eta.$$

Returning to the variables θ and ϕ, we thus obtain, as the final formula,

$$\text{(13)} \qquad \lim \sum \Psi(\theta_i, \phi_l)\, \Delta\theta\, \Delta\phi = \int_\alpha^\beta d\theta \int_{\Phi_0}^{\Phi_1} \Psi(\theta, \phi)\, d\phi.$$

The limits of integration, Φ_0 and Φ_1, are obtained by giving θ a fixed value, $\theta = \theta'$, and then determining the extreme values of ϕ in

S along the line $\theta = \theta'$. The smallest of these values is Φ_0, the largest, Φ_1.

Collecting the results hitherto obtained we are now able to express the volume integral by means of an iterated integral as follows:

$$(14) \qquad \iiint\limits_V f\,dV = \int\limits_{\alpha}^{\beta} d\theta \int\limits_{\Phi_0}^{\Phi_1} d\phi \int\limits_{R_0}^{R_1} f\,r^2 \sin\phi\,dr.$$

The volume integral and the iterated integral are also written in the forms:

$$\iiint\limits_V f\,r^2 \sin\phi\,dr\,d\phi\,d\theta \qquad \text{and} \qquad \int\limits_{\alpha}^{\beta}\int\limits_{\Phi_0}^{\Phi_1}\int\limits_{R_0}^{R_1} f\,r^2 \sin\phi\,dr\,d\phi\,d\theta.$$

We note that, in order to obtain from the integrand f of the volume integral the integrand of the iterated integral, it is necessary to multiply f by $r^2 \sin\phi$, and this is always the first step to take. It is analogous to multiplying the integrand of a double integral by r, when the evaluation of such an integral by means of the iterated integral in polar coordinates is to be employed.

The determination of the limits of integration can be formulated as follows. We have already seen how to find R_0 and R_1. To find Φ_0 and Φ_1 directly from V, without introducing the surface S, give to θ a fixed value, as $\theta = \theta'$, and consider all the points of V which lie in the half-plane $\theta = \theta'$. The smallest ϕ which any one of these points has will be the Φ_0 corresponding to this value of θ; and the largest ϕ will be Φ_1. — Finally, to determine α and β, observe that α is the algebraically smallest value which θ takes on for any point in V, and β, the largest such value.

Example. To find the centre of gravity, G, of a homogeneous hemispherical shell whose radii are a and A.

Let the origin of coordinates be taken at the centre of the spheres, and let the axis of x be the axis of symmetry. Since G lies in this axis, the problem is merely to compute $\bar{x}$. We have:

$$\bar{x} = \frac{\displaystyle\iiint x\,dV}{V}, \qquad V = \frac{2\pi}{3}(A^3 - a^3).$$

Here,
$$f = x = r \sin\phi \cos\theta,$$

and hence the integrand of the iterated integral becomes:

$$f\,r^2 \sin\phi = r^3 \sin^2\phi \cos\theta.$$

The iterated integral itself is

$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} d\theta \int_0^{\pi} d\phi \int_a^A r^3 \sin^2\phi \cos\theta \, dr = \tfrac{1}{4}\pi (A^4 - a^4).$$

Hence $$\bar{x} = \frac{3(a^3 + a^2A + aA^2 + A^3)}{8(a^2 + aA + A^2)}.$$

Checks: (*i*) When a approaches A as its limit, x approaches $\frac{1}{2}A$, and this result agrees with the known position of the centre of gravity of a zone of a sphere. (*ii*) When $a = 0$, $\bar{x} = \frac{3}{8}A$, and we have a solid hemisphere.

EXERCISES

1. Work the foregoing example, using the plane of the base of the shell as the (x, y)-plane, the origin being at the centre and the positive axis of z piercing the shell.

2. The same problem, the plane of the base being in the (x, z)-plane, and the origin at the centre.

3. From the shell just considered a solid is cut by a cone of revolution, co-axial with the shell and of semi-vertical angle α. Find its centre of gravity. Check your answer.

4. Determine the attraction of a material homogeneous shell of the form described in the preceding problem, on a unit particle at the centre of the sphere.

5. Compute the moment of inertia of a homogeneous sphere by triple integration.*

6. Find the centre of gravity of the element of volume represented in Fig. 29.

7. Determine the attraction of a homogeneous solid cone of revolution on a particle situated at its vertex.

8. Think out and work through the evaluation of a volume integral by means of each of the six iterated integrals, of which one is the integral (14). Draw the figure in each case which leads to the double integral. Express in words the rule for determining the limits of integration.

* The term *triple integration* is used to apply both to the *formulation* of a physical quantity as a *volume* integral, and to its *evaluation* by means of an *iterated* integral. A similar remark applies to double integration.

9. Develop the iterated integral by first holding a single coordinate fast (*e.g.* set $r = r_1$) and then obtaining a double integral.

10. Write out the twelve iterated integrals, — six, as three-fold simple integrals, and six as a double integral combined with a simple integral.

11. Apply to Exercise 5 a sufficient number of each type of the iterated integrals considered in the preceding problem to make sure that you understand the rest.

12. A tetrahedron has its faces in the coordinate planes $x = 0$, $y = 0$, and the planes

$$z = x + y, \qquad z = a.$$

Express as an iterated integral of the type (14) the volume integral (2), determining explicitly the limits of integration.

13. The density of a cube is proportional to the distance from its centre. Find its mass.

14. Compute the moment of inertia of the cube of the preceding problem about an axis through the centre parallel to four of the edges.

4. Conclusion; Cylindrical Coordinates. The cylindrical coordinates of a point are defined as in the accompanying figure.* They are a combination of polar coordinates in the (x, y)-plane and the Cartesian z.

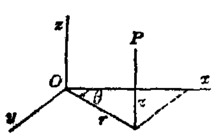


Fig. 31

$$x = r\cos\theta, \qquad y = r\sin\theta, \qquad z = z.$$

The element of volume is shown in Fig. 32. The lengths of the edges adjacent to P, — they meet at right angles there, — are: Δr, $r\Delta\theta$, Δz. Hence the volume, ΔV, of the element differs from $r\,\Delta r\,\Delta\theta\,\Delta z$ by an infinitesimal of higher order, and we have:

$$\lim \frac{\Delta V}{r\,\Delta r\,\Delta\theta\,\Delta z} = 1.$$

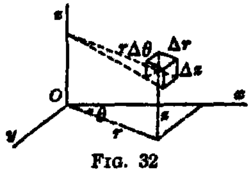


Fig. 32

From Duhamel's Theorem it follows, then, that in taking the limit of the sum (1), § 1, ΔV_k may be replaced by $r_k \Delta r\,\Delta\theta\,\Delta z$, and so, setting

$$f(x, y, z) = F(r, \theta, z),$$

we obtain:

$$\iiint\limits_V f\,dV = \lim_{n=\infty} \sum_{k=1}^{n} F(r_k, \theta_k, z_k)\, r_k\, \Delta r\,\Delta\theta\,\Delta z.$$

* *Analytic Geometry*, p. 587.

This last limit can be computed by iterated integration in a manner precisely similar to that set forth in the case of spherical coordinates. We thus obtain:

$$(15)\qquad \iiint\limits_V f\,dV = \int\limits_a^b dz \int\limits_{\Theta_0}^{\Theta_1} d\theta \int\limits_{R_0}^{R_1} f\,r\,dr,$$

together with similar formulas yielded by adopting a different order of integration.

The above volume integral and the iterated integral are also written in the forms:

$$\iiint\limits_V f\,r\,dr\,d\theta\,dz \quad\text{and}\quad \int\limits_a^b\int\limits_{\Theta_0}^{\Theta_1}\int\limits_{R_0}^{R_1} f\,r\,dr\,d\theta\,dz.$$

Example. To find the attraction of a cylindrical bar on a particle of unit mass situated in its axis.

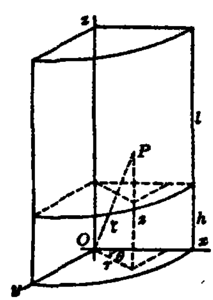


FIG. 33

The attraction is given by the formula (§ 1, Ex. 4):

$$A = \iiint\limits_V \frac{\rho\cos\psi}{\mathfrak{r}^2}\,dV.$$

Here

$$\mathfrak{r}^2 = r^2 + z^2, \qquad \cos\psi = \frac{z}{\mathfrak{r}} = \frac{z}{\sqrt{r^2+z^2}}.$$

Hence

$$A = \rho\int\limits_0^{2\pi} d\theta \int\limits_h^{h+l} dz \int\limits_0^a \frac{zr\,dr}{(r^2+z^2)^{3/2}}.$$

$$(16)\qquad \int\limits_0^a \frac{zr\,dr}{(r^2+z^2)^{3/2}} = -\frac{z}{\sqrt{r^2+z^2}}\Bigg|_0^a = 1 - \frac{z}{\sqrt{a^2+z^2}};$$

$$\int\limits_h^{h+l} dz \int\limits_0^a \frac{zr\,dr}{(r^2+z^2)^{3/2}} = l - \int\limits_h^{h+l} \frac{z\,dz}{\sqrt{a^2+z^2}}$$

$$= l - \sqrt{a^2+(h+l)^2} + \sqrt{a^2+h^2};$$

$$\therefore \qquad A = 2\pi\rho\left[l + \sqrt{a^2+h^2} - \sqrt{a^2+(h+l)^2}\,\right].$$

In the foregoing solution it has been tacitly assumed that h is positive (or zero). If h is negative and, in particular $= -\frac{1}{2}l$, the attraction should clearly be zero. And yet the above formula yields a positive result. What is the trouble?

On scrutinizing the details of the work one finds that equation (16) holds only when $z > 0$. For, if $z < 0$, the value of $\sqrt{r^2+z^2}$, when $r = 0$, is not z, but $-z$. To repeat: it is the *positive* square root that is meant by $\sqrt{\ }$, and not the negative one, and the positive square root of z^2 is here $-z$. Thus (16) must read, when z is negative:

$$(17) \qquad \int_0^a \frac{zr\,dr}{(r^2+z^2)^{3/2}} = -\frac{z}{\sqrt{r^2+z^2}}\Bigg|_0^a = -1 - \frac{z}{\sqrt{a^2+z^2}}.$$

It becomes necessary, therefore, in evaluating the volume integral for the case that $h < 0$, $l + h > 0$, to split the iterated integral into two parts; the first corresponds to the part of the bar above the plane $z = 0$, and this attraction, A_1, is given by the solution in the text, except that the limits of the integral with respect to z are now 0 and $l + h$:

$$A_1 = 2\pi\rho\left[l + h + a - \sqrt{a^2 + (l+h)^2}\right].$$

For the attraction, A_2, of the part below the plane we have:

$$A_2 = 2\pi\rho\left[h - a + \sqrt{a^2+h^2}\right].$$

Hence the resultant attraction, $A = A_1 + A_2$, is:

$$A = 2\pi\rho\left[l + 2h - \sqrt{a^2+(l+h)^2} + \sqrt{a^2+h^2}\right].$$

Check. If in this last formula we set $h = -\frac{1}{2}l$, then $A = 0$, and this result agrees with the physical fact.

EXERCISES

1. Determine the attraction of a straight pipe on a particle situated in its axis.

2. Find the force with which a cone of revolution attracts a particle at its vertex. *Ans.* $2\pi\rho h(1 - \cos\alpha)$.

3. Show that the force with which a piece of a spherical shell cut out by a cone of revolution with its vertex at the centre O attracts a particle at O depends, for a given cone, only on the thickness of the shell.

4. Prove the preceding theorem for any cone.

5. Potential. The *potential* of a system of n particles, of masses $m_1, m_2, \cdots, m_n$, is defined (Chap. V, § 17) as

$$u = \frac{m_1}{r_1} + \frac{m_2}{r_2} + \cdots + \frac{m_n}{r_n},$$

where $r_1, r_2, \cdots, r_n$ denote respectively the distances of the particles from a unit particle situated at a point P.

It is easy to see how this physical conception can be extended to a distribution of matter, continuous throughout a three-dimensional region V of space. Let V be divided into n sub-regions, and let the mass, ΔM_k, of the k-th of these be concentrated at one of its points. Then, when n is large and the longest diameter of any sub-region is small, the sum

$$\frac{\Delta M_1}{r_1} + \frac{\Delta M_2}{r_2} + \cdots + \frac{\Delta M_n}{r_n},$$

i.e. the potential of the n particles, appeals to our physical intuition as representing approximately what we should understand by the potential of the continuous distribution; and we should expect the approximation to increase in accuracy and approach as its limit the potential, u, in question. Thus

$$u = \lim_{n=\infty} \sum_{k=1}^{n} \frac{\Delta M_k}{r_k}.$$

Since $\Delta M_k = \bar{\rho}_k \Delta V_k$, where $\bar{\rho}_k$ denotes the average density of the k-th piece, the above limit is equal to the volume integral:

(1) $$u = \iiint\limits_V \frac{\rho\, dV}{r}.$$

Similarly, if matter be distributed continuously over a curved surface, the potential at a point P not on the surface is

(2) $$u = \iint\limits_S \frac{\sigma\, dS}{r},$$

where σ denotes the density of the distribution; r, the distance from P to a variable point on the surface, and the integral is the surface integral extended over the surface.

The foregoing definitions apply, with the obvious modifications in form, to continuous distributions of electricity.

Example. **To find the potential at an interior point of a homogeneous shell bounded by concentric spheres of radii a and A.**

Let the axes be so chosen that P lies on the positive axis of z, the origin being at the centre, and let its distance from the centre of the shell be h. On introducing spherical coordinates, we have:

$$\mathfrak{r}^2 = r^2 + h^2 - 2\,hr\cos\phi,$$

where $\mathfrak{r}$ denotes the denominator of the integrand in (1); *i.e.* the distance from P to a variable point (r, θ, ϕ) of the distribution. Hence

$$u = \rho\int_0^{2\pi} d\theta\int_a^A dr\int_0^\pi \frac{r^2\sin\phi\, d\phi}{\sqrt{r^2 + h^2 - 2\,hr\cos\phi}}.$$

$$\int_0^\pi \frac{r^2\sin\phi\, d\phi}{\sqrt{r^2+h^2-2\,hr\cos\phi}} = \frac{r}{h}\sqrt{r^2+h^2-2\,hr\cos\phi}\,\Bigg|_0^\pi = 2\,r.$$

Thus

$$u = 2\,\pi\rho(A^2 - a^2),$$

and we are led to the result: *The potential is constant within the above spherical shell.*

Furthermore, since the force which a distribution of matter exerts in any direction is proportional to the directional derivative of the potential function, it follows that the force is nil at each interior point of the shell.

Remark. The potential of a homogeneous sphere of radius R at its centre is

$$2\,\pi\rho R^2 = \frac{3}{2}\frac{M}{R}.$$

EXERCISES

1. Show that the potential of the shell of the Example at any exterior point is the same as that of a particle of like mass, situated at the centre of the shell.

2.* Obtain the potential of a uniform spherical lamina at any interior point (*i*) by evaluating the appropriate surface integral, taken over the sphere; (*ii*) by allowing a to approach A in the result of the text, the mass of the shell being held fast. *Ans.* M/R.

3. The same problem for an exterior point. *Ans.* M/r.

4. Show that a homogeneous sphere attracts a particle outside it as if all its mass were concentrated at its centre. Give the solution first by means of the results obtained in this paragraph. Secondly, compute the attraction directly, as an exercise in triple integration.

* The results in Exs. 2 and 3 are important in the case of a uniform distribution of electricity over a spherical surface.

5. Show that a homogeneous sphere attracts a particle situated in its interior with a force proportional to the distance from the centre.

Suggestion. Pass a concentric spherical surface through the particle and consider the two distributions into which the sphere is thus divided.

6. If the density of a sphere is continuous, and if it is constant over any concentric spherical surface; *i.e.* if it depends only on the distance from the centre, show that the potential at an exterior point is the same as if all the mass were concentrated at the centre.

7. Refine the physical hypothesis in a way analogous to that followed in the case of moments of inertia, fluid pressures, attractions, etc., and thus, with the aid of Duhamel's Theorem, deduce the formulas of the text, (1) and (2), for the potential.

EXERCISES ON CHAPTER IV

1. Determine the attraction of a bar, of rectangular cross-section, on an exterior particle situated in its axis.

2. Write down the five equivalent forms of the integral

$$\int_0^a dy \int_0^y dx \int_0^x f(x, y, z)\, dz,$$

obtained by changing the order of the integrations.

3. Two spheres are tangent to each other internally, and also to the (x, y)-plane at the origin. Denoting the space included between the spheres by V, express the volume integral

$$\iiint_V f\, dV$$

by means of iterated integrals in Cartesian coordinates.

4. Express the iterated integral

$$\int_0^a dx \int_0^{\frac{b}{a}\sqrt{a^2-x^2}} dy \int_{x+y}^{2+4x+5y} f\, dz$$

as a volume integral, and state throughout what region of space the latter is to be extended.

5. The same for

$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos\theta\, d\theta \int_0^{\frac{\pi}{4}} \sin\phi\, d\phi \int_{2b\cos\phi}^{2a\cos\phi} dr.$$

6. The temperature within a spherical shell is inversely proportional to the distance from the centre, and has the value T_0 on the inner surface. Given that the quantity of heat required to raise any piece of the shell from one uniform temperature to another is proportional jointly to the volume of the piece and the rise in temperature, and that C units of heat are required to raise the temperature of a cubic unit of the shell by one degree, find how much heat the shell will give out in cooling to the temperature 0°.

Ans. $2\pi C T_0 a(b^2 - a^2)$

7. The interior of an iron pipe is kept at 100° C. and the exterior at 15°. The length of the inner radius of the pipe is 2 cm., that of the outer radius, 3 cm. The temperature at any interior point is given by the formula:

$$T = \alpha \log r + \beta,$$

where r is the distance from the axis and the constants α, β are to be determined from the above data. Taking the specific heat of iron as .11, and its specific gravity as 7.8, how much heat will a segment of the pipe 30 cm. long give out in cooling to 0°?

Ans. 21,000 calories.

8. Show that the attraction of a homogeneous spherical segment of one base, on a particle situated at its vertex, is

$$2\pi\rho h\left\{1 - \frac{1}{3}\sqrt{\frac{2h}{a}}\right\},$$

where a denotes the radius of the sphere and h, the altitude of the segment.

9. Show that the segment of the preceding question attracts a particle situated at the centre of the base with a force

$$\frac{2\pi\rho h}{3(a-h)^2}[3a^2 - 3ah + h^2 - (2a-h)^{\frac{3}{2}}h^{\frac{1}{2}}].$$

10. "Show that the attraction of a homogeneous segment of one base of a paraboloid of revolution, on a particle situated at the focus, is

$$4\pi\rho a \log \frac{a+b}{ea},$$

where a denotes the distance from the vertex to the focus, and b is the altitude of the segment."

Prove that this proposition is true when $b \geqq a$, and correct it when $b < a$.

11. Find the attraction of the solid of the preceding problem on a particle situated at the vertex.

12. Show that, in the case of a homogeneous oblique cone whose base is any plane figure, the attraction at the vertex due to any frustum is proportional to the thickness of the frustum.

13. A homogeneous hemisphere attracts a particle situated in the rim of its base. Show that the component perpendicular to the base is $\frac{2}{3}\pi\rho a$.

14. The Great Pyramid is 481 ft. high, and the base is 756 ft. on a side. If it were homogeneous and of density equal to the average density of the earth, — namely, 5.6 times that of water, — find the force with which it would attract a mass of one ton situated at its vertex. (For the value of the gravitational constant, cf. the *Introduction to the Calculus*, p. 334.)

CHAPTER V

PARTIAL DIFFERENTIATION

1. Functions of Several Variables. Limits and Continuity.* Consider a region, S, of the (x, y)-plane. To each point (x, y) of S let there be assigned a definite number, u, according to any specific rule. Then u is called a *function* of the independent variables (x, y) in the region S, and is denoted, for example, by the notation

(1) $$u = f(x, y).$$

Similarly, we may consider a region V of three-dimensional space and assign a number, u, to each of its points. Then u is a function of the three independent variables which are the coordinates of a point in the region V:

(2) $$u = f(x, y, z).$$

When the number of independent variables exceeds three, our geometric intuition fails to provide us with a picture of a region in four-, five-, or n-dimensional space. It is convenient to speak of such regions by analogy with space of two or three dimensions; but the foundation for the definition of such a region must be sought in an analytic formulation. Thus we might consider those *points* † (x, y, z, t) of four-dimensional space whose *coordinates* satisfy the relation

$$x^2 + y^2 + z^2 + t^2 < 1,$$

and call this region, by way of analogy, a *four-dimensional hypersphere.*

* For the beginner, this first paragraph should be regarded as primarily *descriptive.* He should read it thoughtfully for the ideas it suggests; but it should not be made a task, like certain courses in History, which is tested by a premature examination. Its object, at this stage, is *cultural,* — to give the student background, to acquaint him with the great ideas that underlie this domain of analysis, and also to supply him with such information as he needs in the immediate future. As he proceeds with the later paragraphs, he will do well frequently to recur to these pages, for they will mean more to him as his knowledge increases and his imagination develops. For a thorough understanding of the subject of this chapter, this paragraph is of the highest importance.

† Meaning thereby nothing more or less than the quadruple of values (x, y, z, t), *i.e.* this *mark* itself.

Fortunately that which is novel in functions of several variables can, for the most part, be set forth by examining functions of two and three independent variables. In the first case the function, (1), can be represented geometrically by a surface.* The function is said to be *continuous* at a point (x_0, y_0) of S when the surface is continuous at the corresponding point. Let us seek an arithmetic definition of continuity which can be applied to functions of any number of variables.

Limits. Let $f(x, y)$ be defined at all points of the neighborhood of a point (a, b) with the exception of this one point itself.† Then $f(x, y)$ is said to approach a *limit*, A:

$$\lim_{(x,y)\doteq(a,b)} f(x, y) = A,$$

provided $f(x, y)$ satisfies the following condition. Let ϵ be a positive number, chosen at pleasure, but then held fast. Then it shall be possible to find a second positive number δ, having the property that the relation

$$|f(x, y) - A| < \epsilon$$

shall hold for every point, except (a, b), whose coordinates satisfy the relations ‡

$$|x - a| < \delta, \qquad |y - b| < \delta.$$

Continuity. A function $f(x, y)$ is said to be *continuous* at a point (x_0, y_0) if it is defined at every point in the neighborhood of (x_0, y_0) and

$$\lim_{(x,y)\doteq(x_0,y_0)} f(x, y) = f(x_0, y_0).$$

* There is, really, a subtle question here involved, namely, that of whether the function is the dependent variable, u, which is represented by the *ordinate*, or whether it is not rather the locus of the triples (u, x, y) which is represented by the *surface*. As a matter of fact, the word *function* is used in both senses; but it is the former sense in which it most frequently appears in what follows.

† The function may be defined at the point (a, b), too. But, if it is, this fact is wholly irrelevant in the definition we are engaged in setting forth.

‡ Geometrically the definition can be illustrated as follows. Represent the function by a surface as described above, $u = f(x, y)$. Draw the horizontal planes

(*i*) $$u = A + \epsilon, \qquad u = A - \epsilon$$

and erect the vertical planes

(*ii*) $$x = a + \delta, \qquad x = a - \delta, \qquad y = b + \delta, \qquad y = b - \delta.$$

These six planes enclose a parallelepiped. And now, to say that $f(x, y)$ approaches A as its limit is merely to say that those points (u, x, y) of the locus $u = f(x, y)$ for which (x, y) lies in the square (*ii*) of the (x, y)-plane — the centre (a, b) being omitted — are all contained within this parallelepiped.

In case (x_0, y_0) lies on the boundary of the region, only such points (x, y) come into consideration as lie in the region.

These definitions extend to functions of any number of variables.

Infinitesimals. Let

$$\zeta = f(\alpha, \beta)$$

be a function which is defined throughout the neighborhood of the origin, $(\alpha, \beta) = (0, 0)$, with the possible exception of this one point itself; and let

$$\lim_{(\alpha, \beta) \doteq (0, 0)} \zeta = 0.$$

Then ζ is called an *infinitesimal.* The independent variables (α, β) are called the *principal infinitesimals.*

The concept of *order,* — same order, first order, second order, etc., — does not admit of immediate or useful extension to the infinitesimals under consideration, except in the following case. We define ζ to be an infinitesimal of *higher order* provided

$$\lim_{(\alpha, \beta) \doteq (0, 0)} \frac{\zeta}{\sqrt{\alpha^2 + \beta^2}} = 0.$$

We might equally well lay down the definition in the form *

$$\lim_{(\alpha, \beta) \doteq (0, 0)} \frac{\zeta}{|\alpha| + |\beta|} = 0.$$

Thus

$$\zeta = \alpha^2 + \alpha\beta + \beta^2$$

is an infinitesimal of higher order, (α, β) being the principal infinitesimals.

2. Law of the Mean for Functions of a Single Variable. Let $f(x)$ be a function which is continuous throughout the interval $a \leqq x \leqq b$, and let it have a derivative, $df/dx = f'(x)$, at every interior point of the interval. Draw the graph, and let LM be the secant connecting its extremities. Then there will be at least one point of the graph at which the tangent to the graph is parallel to the secant LM.

The truth of this statement is evident intuitively. For, consider the distance, PQ, from a point P of the curve to the secant, measured along an ordinate. This distance (taken algebraically) will have either a maximum or a minimum value, and at such a point the

* The latter form has the advantage that it can be extended to complex quantities, whereas the former form then breaks down.

tangent is evidently parallel to the secant. Now the slope of the secant is

$$\tan \angle NLM = \frac{f(b) - f(a)}{b - a},$$

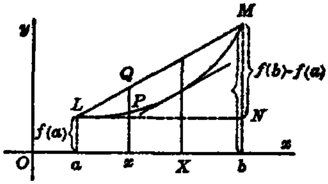


FIG. 34

and the slope of the curve at $x = X$ is $f'(X)$. Hence

$$\frac{f(b) - f(a)}{b - a} = f'(X),$$

(A) $$f(b) - f(a) = (b - a)f'(X), \qquad a < X < b.$$

If we set $b - a = h$, then $b = a + h$, and we may write X in the form:

$$X = a + \theta h,$$

where θ is some number lying between 0 and 1.* Equation (*A*) can now be written in the equivalent form:

(B) $$f(a + h) = f(a) + hf'(a + \theta h), \qquad 0 < \theta < 1.$$

The theorem contained in either of the equations, (*A*) or (*B*), is known as the Law of the Mean in the Differential Calculus.† In the form (*B*) it is identical with Taylor's Theorem with the Remainder for the simplest case.

In (*A*), a and b can be interchanged, and in (*B*), h can be negative.

An analytical proof of the Law of the Mean can be given as follows. Form the function

$$\phi(x) = \frac{f(b) - f(a)}{b - a}(x - a) - [f(x) - f(a)].$$

This function satisfies all the conditions of Rolle's Theorem, *Introduction*, p. 430, and hence its derivative,

$$\phi'(x) = \frac{f(b) - f(a)}{b - a} - f'(x),$$

must vanish for a value $x = X$ between a and b:

$$\frac{f(b) - f(a)}{b - a} - f'(X) = 0, \qquad a < X < b.$$

Thus the theorem is proven.

* We may think of the second term, θh, as representing that portion of the interval $b - a = h$ which must be added to the segment a to take us to X.

† The Law of the Mean in the Integral Calculus was obtained in the *Introduction to the Calculus*, p. 343.

This proof merely puts into analytic form the geometric proof first given, for the function $\phi(x)$ here employed is precisely the distance PQ.*

EXERCISES

1. Show that

$$\frac{h}{1+h} < \log(1+h) < h, \qquad 0 < h.$$

2. Show that

$$h < \log\frac{1}{1-h} < \frac{h}{1-h}, \qquad 0 < h < 1.$$

3. Show that

$$1 + \frac{x}{2\sqrt{1+x}} < \sqrt{1+x} < 1 + \tfrac{1}{2}x, \quad -1 < x, \quad x \neq 0.$$

3. The Fundamental Lemma. We have already defined partial derivatives.† Let u be a function of several independent variables, as x, y, z:

$$u = f(x, y, z),$$

and let all the variables but x be held fast. Then u becomes a function of x alone, and its derivative is denoted by ‡

$$\frac{\partial u}{\partial x} \quad \text{or} \quad f_1(x, y, z) \quad \text{or} \quad f_x(x, y, z).$$

Similarly, when y alone is allowed to vary, we have

$$\frac{\partial u}{\partial y} \quad \text{or} \quad f_2(x, y, z) \quad \text{or} \quad f_y(x, y, z), \quad \text{etc.}$$

There are as many partial derivatives of the first order as there are independent variables.

Furthermore, we write

$$\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial y}\right) = \frac{\partial^2 u}{\partial x\, \partial y} = f_{12}(x, y) \quad \text{or} \quad f_{xy}(x, y),$$

$$\frac{\partial^2 u}{\partial x^2} = f_{11}(x, y) \quad \text{or} \quad f_{xx}(x, y), \quad \text{etc.}$$

* The underlying importance of the analytic proof is due to the fact that this proof rests on the most elementary considerations of analysis, as distinguished from geometry, — namely, on the theorems about continuity and the definition of a derivative. A detailed study of these, however, belongs to a later stage.

† *Introduction to the Calculus*, Chap. XV, § 2.

‡ It is not possible to consider the expression $\partial u/\partial x$ as the ratio of two infinitesimals; cf. § 5. The notation must be taken as a whole, which expresses the partial derivative.

The theory of partial differentiation is based on a lemma which we proceed to deduce. For convenience let the number of independent variables be two:

(1) $$u = f(x, y),$$

and let this function, together with its partial derivatives of the first order, be continuous throughout a region, S.

Consider an arbitrary point, (x_0, y_0), and a second point,

$$(x_0 + \Delta x, y_0 + \Delta y).$$

Denote the corresponding increment in the function by Δu:

(2) $$\Delta u = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0).$$

This last expression we now transform by adding and subtracting the same quantity, $f(x_0, y_0 + \Delta y)$:

(3) $$\Delta u = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0 + \Delta y) + f(x_0, y_0 + \Delta y) - f(x_0, y_0).$$

To the first of these differences we apply the Law of the Mean, § 2, thinking of $f(x, y_0 + \Delta y)$ as a function of x alone, and letting x range through the interval $x_0 \leqq x \leqq x_0 + \Delta x$. The derivative of this function is $f_1(x, y_0 + \Delta y)$. Thus the Law of the Mean yields the result:

$$f(x_0+\Delta x, y_0+\Delta y) - f(x_0, y_0+\Delta y) = \Delta x f_1(x_0+\theta\,\Delta x, y_0+\Delta y), \quad 0<\theta<1.$$

To transform the second line in the expression for Δu, consider the function of y alone, $f(x_0, y)$, in the interval $y_0 \leqq y \leqq y_0 + \Delta y$. The derivative of this function is $f_2(x_0, y)$, and thus the Law of the Mean gives:

$$f(x_0, y_0 + \Delta y) - f(x_0, y_0) = \Delta y\, f_2(x_0, y_0 + \theta'\Delta y), \quad 0 < \theta' < 1.$$

Hence we have the relation:

(4) $$\Delta u = f_1(x_0 + \theta\,\Delta x, y_0 + \Delta y)\,\Delta x + f_2(x_0, y_0 + \theta'\,\Delta y)\,\Delta y.$$

It is assumed that all points of the rectangle whose vertices lie in the four points $(x_0 \pm \Delta x, y_0 \pm \Delta y)$ lie in the region S.

By hypothesis, the partial derivatives,

$$\frac{\partial u}{\partial x} = f_1(x, y), \qquad \frac{\partial u}{\partial y} = f_2(x, y)$$

are continuous functions. If, then, we set

$$f_1(x_0+\theta\,\Delta x, y_0+\Delta y) = f_1(x_0, y_0) + \epsilon, \qquad f_2(x_0, y_0+\theta'\,\Delta y) = f_2(x_0, y_0) + \eta,$$

both ϵ and η will be infinitesimals:

$$\lim_{(\Delta x,\,\Delta y)\doteq(0,\,0)} \epsilon = 0, \qquad \lim_{(\Delta x,\,\Delta y)\doteq(0,\,0)} \eta = 0.$$

On substituting these values in (4), we have:

(5) $$\Delta u = f_1(x_0, y_0)\,\Delta x + f_2(x_0, y_0)\,\Delta y + \epsilon\,\Delta x + \eta\,\Delta y,$$

or, on dropping the subscripts,

(I) $$\Delta u = \frac{\partial u}{\partial x}\Delta x + \frac{\partial u}{\partial y}\Delta y + \epsilon\,\Delta x + \eta\,\Delta y,$$

where (x, y) is an arbitrary point; Δx and Δy are any two increments, subject merely to the condition that the rectangle with vertices $(x \pm \Delta x,\ y \pm \Delta y)$ lie in S; and ϵ, η are infinitesimals whenever Δx and Δy are infinitesimals.

If u is a function of three variables, x, y, z, the equation takes the form

(I′) $$\Delta u = \frac{\partial u}{\partial x}\Delta x + \frac{\partial u}{\partial y}\Delta y + \frac{\partial u}{\partial z}\Delta z + \epsilon\,\Delta x + \eta\,\Delta y + \zeta\,\Delta z,$$

where ϵ, η, ζ are infinitesimal when Δx, Δy, Δz are infinitesimal; and similarly for any number of variables.

4. Change of Variables. If u comes to us as a function of the variables (x, y),

(1) $$u = f(x, y),$$

and we make x and y depend on new variables, (r, s):

(2) $$x = \phi(r, s), \qquad y = \psi(r, s),$$

then u becomes a function of (r, s). The derivative of u with respect to r is expressed by means of the following formula:

(A) $$\frac{\partial u}{\partial r} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial r}.$$

To prove this statement, choose an arbitrary point (r_0, s_0), give r an increment Δr, and denote the corresponding increments in u, x, and y respectively by $\Delta_r u$, $\Delta_r x$, and $\Delta_r y$. Then, by the lemma of § 3,

$$\Delta_r u = \frac{\partial u}{\partial x}\Delta_r x + \frac{\partial u}{\partial y}\Delta_r y + \epsilon\,\Delta_r x + \eta\,\Delta_r y.$$

On dividing through by Δr and allowing Δr to approach 0 as its limit, we have:

$$\lim_{\Delta r \doteq 0}\frac{\Delta_r u}{\Delta r} = \lim_{\Delta r \doteq 0}\frac{\partial u}{\partial x}\frac{\Delta_r x}{\Delta r} + \lim_{\Delta r \doteq 0}\frac{\partial u}{\partial y}\frac{\Delta_r y}{\Delta r}.$$

Here, $\partial u/\partial x$ and $\partial u/\partial y$ are constants, being the values of these derivatives at the point (x_0, y_0) which corresponds to (r_0, s_0). Hence they may be taken outside the limit sign. Moreover,

$$\lim_{\Delta r \doteq 0} \frac{\Delta_r u}{\Delta r} = \frac{\partial u}{\partial r}, \qquad \lim_{\Delta r \doteq 0} \frac{\Delta_r x}{\Delta r} = \frac{\partial x}{\partial r}, \qquad \lim_{\Delta r \doteq 0} \frac{\Delta_r y}{\Delta r} = \frac{\partial y}{\partial r}.$$

Thus the truth of (A) is established.

It is assumed that ϕ and ψ are continuous, together with their partial derivatives of the first order, throughout a region Σ of the (r, s)-plane; that, moreover, f is continuous, together with its partial derivatives of the first order, throughout a region S of the (x, y)-plane; and that, finally, the points (x, y) which correspond by (2) to points (r, s) of Σ lie in S.

The number of variables in the two classes, — (x, y) on the one hand, and (r, s) on the other, — need not be the same; the number in each class is arbitrary. Thus the variables of the first class might be x, y, z, and those of the second class, the single variable, t. The derivative of u with respect to t would then be a total derivative, and we should have:

$$\frac{du}{dt} = \frac{\partial u}{\partial x}\frac{dx}{dt} + \frac{\partial u}{\partial y}\frac{dy}{dt} + \frac{\partial u}{\partial z}\frac{dz}{dt}. \tag{3}$$

Again, there may be but a single variable, x, in the first class, and several, $(r, s, \cdots)$ in the second class. Here,

$$\frac{\partial u}{\partial r} = \frac{du}{dx}\frac{\partial x}{\partial r}, \qquad \frac{\partial u}{\partial s} = \frac{du}{dx}\frac{\partial x}{\partial s}, \qquad \text{etc.} \tag{4}$$

If there is but one variable in each class, the case reduces to that of the *Introduction*, Chap. II, § 8, p. 35; and on the other hand the present theorem is a generalization of that one.

The result can be formulated as follows.

THEOREM. *If u be a function of the variables $x, y, z, \cdots$:*

$$u = f(x, y, z, \cdots),$$

continuous, together with its partial derivatives of the first order, throughout a region S of the $(x, y, z, \cdots)$-space; and if each of the arguments $x, y, z, \cdots$ be set equal to a function of the variables $r, s \cdots$:

$$x = \phi(r, s, \cdots), \qquad y = \psi(r, s, \cdots), \qquad z = \omega(r, s, \cdots), \cdots,$$

where $\phi, \psi, \omega, \cdots$ are continuous, together with their partial derivatives of the first order, throughout a region Σ of the $(r, s, \cdots)$-space, and where, *moreover, each point $(r, s, \cdots)$ of Σ leads to a point $(x, y, z, \cdots)$ of S; then*

$$(A) \qquad \frac{\partial u}{\partial r}=\frac{\partial u}{\partial x}\frac{\partial x}{\partial r}+\frac{\partial u}{\partial y}\frac{\partial y}{\partial r}+\frac{\partial u}{\partial z}\frac{\partial z}{\partial r}+\cdots,$$

with similar equations for $\partial u/\partial s, \cdots$.

Example 1. Let

$$u=e^{xy},$$

$$x=\log\sqrt{r^2+s^2}, \qquad y=\tan^{-1}\frac{s}{r}.$$

Then

$$\frac{\partial u}{\partial x}=ye^{xy}, \qquad \frac{\partial u}{\partial y}=xe^{xy},$$

$$\frac{\partial x}{\partial r}=\frac{r}{r^2+s^2}, \qquad \frac{\partial y}{\partial r}=\frac{-s}{r^2+s^2},$$

and hence

$$\frac{\partial u}{\partial r}=\frac{ry-sx}{r^2+s^2}e^{xy},$$

from which expression x and y can be eliminated if desired.

Example 2. If

$$u=f(x+a, y+b),$$

show that

$$\frac{\partial u}{\partial x}=\frac{\partial u}{\partial a} \quad \text{and} \quad \frac{\partial u}{\partial y}=\frac{\partial u}{\partial b}.$$

Here, u is not an arbitrary function of the four variables x, y, a, b, but depends on these variables only as they enter through their sums, $x+a$ and $y+b$. In other words, u is any function of *two* variables, X and Y, continuous together with its first derivatives, and these variables in turn are set equal to the above sums:

$$u=f(X, Y),$$

$$X=x+a, \qquad Y=y+b.$$

The derivatives of u with respect to the variables of the second class, (x, y, a, b), can be computed by the theorem of this paragraph:

$$\frac{\partial u}{\partial x}=\frac{\partial u}{\partial X}\frac{\partial X}{\partial x}+\frac{\partial u}{\partial Y}\frac{\partial Y}{\partial x}$$

$$=f_1\cdot 1 + f_2\cdot 0.$$

Similarly,

$$\frac{\partial u}{\partial a}=f_1,$$

and hence

$$\frac{\partial u}{\partial x}=\frac{\partial u}{\partial a}.$$

In like manner the second equation is established.

Remark. In applying the theorem for the change of variables, which is embodied in Formula (A), the student must make clear to himself at the outset that he has to do with *two* classes of independent variables, — the variables of the *first class* corresponding to the (x, y) of the text, and those of the *second class* corresponding to the (r, s) of the text. In the terms on the right-hand side of (A), the first factor is each time a derivative with respect to a variable of the *first class*, and there are as many terms as there are variables of this class. The second factor is the derivative of one of the functions (2) with respect to the particular variable of the *second class*, which has been singled out; and there are as many different equations (A) as there are variables of the second class, for the complete solution of the problem consists in finding not only $\partial u/\partial r$, but also $\partial u/\partial s$, etc.

EXERCISES

1. If
$$u = x^2 - y^2$$
and
$$\begin{cases} x = 2r - 3s + 7, \\ y = -r + 8s - 9, \end{cases}$$
find $\dfrac{\partial u}{\partial r}$. *Ans.* $\dfrac{\partial u}{\partial r} = 4x + 2y$.

2. In the preceding question, find $\dfrac{\partial u}{\partial s}$.

3. If
$$u = xy^z$$
and
$$x = a\cos\theta, \qquad y = a\sin\theta, \qquad z = b\theta,$$
find $\dfrac{du}{d\theta}$.

4. If
$$u = \frac{x+y}{1-xy}$$
and
$$x = \tan(2r - s^2), \qquad y = \cot(r^2 s),$$
find $\dfrac{\partial u}{\partial r}$ and $\dfrac{\partial u}{\partial s}$.

5. If
$$u = f(x, y, z)$$
and
$$\left.\begin{aligned} x &= ax' + by' + cz', \\ y &= a'x' + b'y' + c'z', \\ z &= a''x' + b''y' + c''z', \end{aligned}\right\}$$
show that
$$\frac{\partial u}{\partial x'} = a\frac{\partial u}{\partial x} + a'\frac{\partial u}{\partial y} + a''\frac{\partial u}{\partial z},$$
and find $\dfrac{\partial u}{\partial y'}$ and $\dfrac{\partial u}{\partial z'}$.

6. If
$$x = r\cos\phi, \qquad y = r\sin\phi,$$
show that
$$\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 = \left(\frac{\partial u}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial u}{\partial \phi}\right)^2.$$

Suggestion. Compute first $\dfrac{\partial u}{\partial r}$ and $\dfrac{\partial u}{\partial \phi}$ in terms of $\dfrac{\partial u}{\partial x}$ and $\dfrac{\partial u}{\partial y}$.

7. If
$$u = f(xy),$$
show that
$$x\frac{\partial u}{\partial x} = y\frac{\partial u}{\partial y}.$$

8. If
$$u = f\left(\frac{y}{x}\right),$$
show that
$$x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y} = 0.$$

5. The Total Differential. The Fundamental Lemma, Formula (I) of § 3:
$$\Delta u = \underbrace{\frac{\partial u}{\partial x}\Delta x + \frac{\partial u}{\partial y}\Delta y} + \underbrace{\epsilon\,\Delta x + \eta\,\Delta y},$$

affords an *analysis*, or *breaking up*, of the increment, Δu, into two parts, each of which is simple for its own peculiar reason. The first two terms form a function of Δx and Δy of the simplest imaginable type,—a *linear* function, for $\partial u/\partial x$ and $\partial u/\partial y$ do not depend on Δx and Δy. It is natural to define these terms as the *principal part* of the infinitesimal Δu. The remaining terms constitute an infinitesimal of higher order. For,
$$\frac{\epsilon\,\Delta x + \eta\,\Delta y}{\sqrt{\Delta x^2 + \Delta y^2}} = \epsilon\frac{\Delta x}{\sqrt{\Delta x^2 + \Delta y^2}} + \eta\frac{\Delta y}{\sqrt{\Delta x^2 + \Delta y^2}},$$
and each of the fractions on the right-hand side is numerically less than, or at most equal to, unity.

Hence
$$\lim_{(\Delta x,\,\Delta y)\doteq(0,\,0)} \frac{\epsilon\,\Delta x + \eta\,\Delta y}{\sqrt{\Delta x^2 + \Delta y^2}} = 0.$$

We define the *differential* of u as the principal part of Δu, and write:

(1)
$$du = \frac{\partial u}{\partial x}\Delta x + \frac{\partial u}{\partial y}\Delta y,$$

noting well that the independent variables of the function u are x and y.

Since the definition holds for all functions u, we may in particular set $u = x$. It follows, then, that

$$dx = \frac{\partial x}{\partial x}\Delta x + \frac{\partial x}{\partial y}\Delta y$$

$$= 1 \cdot \Delta x + 0 \cdot \Delta y,$$

or

$$dx = \Delta x.$$

Similarly, on setting $u = y$, we infer that

$$dy = \Delta y.$$

On substituting these values in (1) we have:

$$(B) \qquad du = \frac{\partial u}{\partial x}dx + \frac{\partial u}{\partial y}dy.$$

Thus far, the independent variables have been x and y, and the infinitesimals dx, dy, being equal respectively to the increments Δx, Δy, are independent, or principal, infinitesimals. If we introduce new variables as in § 4, setting

$$(2) \qquad x = \phi(r, s), \qquad y = \psi(r, s),$$

then dx and Δx will in general no longer be equal, and the same is true of dy and Δy. Hence equations (1) and (B) cannot in general both be true, and there is no *a priori* reason to suppose that either will be.

FUNDAMENTAL THEOREM: *The equation*

$$(B) \qquad du = \frac{\partial u}{\partial x}dx + \frac{\partial u}{\partial y}dy$$

holds, no matter what the independent variables be.

Proof. When r and s are the independent variables, we have by definition:

$$(3) \qquad \begin{cases} dx = \dfrac{\partial x}{\partial r}\Delta r + \dfrac{\partial x}{\partial s}\Delta s, \\[2ex] dy = \dfrac{\partial y}{\partial r}\Delta r + \dfrac{\partial y}{\partial s}\Delta s, \end{cases}$$

where x and y are respectively the functions $\phi(r, s)$ and $\psi(r, s)$ of (2). On the other hand, by definition,

$$(4) \qquad du = \frac{\partial u}{\partial r}\Delta r + \frac{\partial u}{\partial s}\Delta s.$$

Multiply the first of equations (3) through by $\partial u/\partial x$, the second by $\partial u/\partial y$, and add. Thus

(5)
$$\frac{\partial u}{\partial x}dx+\frac{\partial u}{\partial y}dy=$$
$$\left[\frac{\partial u}{\partial x}\frac{\partial x}{\partial r}+\frac{\partial u}{\partial y}\frac{\partial y}{\partial r}\right]\Delta r+\left[\frac{\partial u}{\partial x}\frac{\partial x}{\partial s}+\frac{\partial u}{\partial y}\frac{\partial y}{\partial s}\right]\Delta s.$$

But the brackets are equal by Theorem (A) of § 4 respectively to $\partial u/\partial r$ and $\partial u/\partial s$, and hence the right-hand side of (5) reduces to the right-hand side of (4). Hence the left-hand sides are equal, q. e. d.

The number of variables in each class is arbitrary. In particular, if u depends on x, y, z, we have:

$$du=\frac{\partial u}{\partial x}dx+\frac{\partial u}{\partial y}dy+\frac{\partial u}{\partial z}dz.$$

And again, if there is only one variable, as x, we have the equation of the Theorem on p. 93 of the *Introduction*:

$$du=\frac{du}{dx}dx.$$

But this equation now holds not merely when x depends on a single variable, as t, but when x is a function of any number of variables, as r, s, $\cdots$.

It is possible to look on the individual terms in the right-hand side of equation (B) as the principal parts of the partial increments in the function u, due to varying one argument at a time:

$$\Delta_x u=f(x+\Delta x,\,y)-f(x,\,y),$$
$$d_x u=\frac{\partial u}{\partial x}dx,\qquad \text{etc.,}$$

and to write:

$$du=d_x u+d_y u.$$

From this point of view, du is spoken of as the *total differential* of u, and it appears as equal to the sum of all the *partial differentials*, $d_x u$, $d_y u$. But these partial differentials are of little use in practice, for it is not possible to pick to pieces a partial derivative, $\partial u/\partial x$, and regard it as the quotient of an infinitesimal, ∂u, by a second infinitesimal, ∂x.*

* The equation we should like to write from this point of view, namely,

$$\partial u=\frac{\partial u}{\partial x}\partial x+\frac{\partial u}{\partial y}\partial y,$$

The geometric representation of the differential in the case of functions of two independent variables has already been pointed out; *Introduction*, p. 444

It is readily shown that the general theorems relating to the differentials of functions of a single variable:

$$d(cu) = c\,du$$
$$d(u+v) = du + dv,$$
$$d(uv) = u\,dv + v\,du,$$
$$d\left(\frac{u}{v}\right) = \frac{v\,du - u\,dv}{v^2},$$

hold for functions of several variables. Moreover, the differential of a constant, considered as a function of several variables, is 0:

$$dc = 0.$$

Remark. The student may find himself confronted by a subtle difficulty in the theorem that "the differential of an independent variable is equal to the increment of that variable." For, differentials have been defined only for functions, *i.e.* dependent variables. There are two ways out: (*i*) define the differential of an independent variable as equal to the increment of that variable; (*ii*) refrain altogether from defining the differential of an independent variable and consider dx, when x is an independent variable, to be the differential of the function, $u = x$.

The second alternative corresponds precisely to what is done, in an analogous case in the Calculus of Variations, in defining the variations δy, δz, $\cdots$ of the independent functions, y, z, $\cdots$.

6. Continuation. Applications. By means of differentials it is possible to compute the partial derivative of a function when a change of variables has taken place. Let us treat the two Examples of § 4 by the new method.

Example 1.
$$u = e^{xy},$$
$$x = \log\sqrt{r^2+s^2}, \qquad y = \tan^{-1}\frac{s}{r}.$$

---

leads to the absurdity that

$$1 = 2.$$

We refrain, therefore, once for all from undertaking to give to ∂x and ∂u any independent meanings, and regard the notation $\dfrac{\partial u}{\partial x}$ as one homogeneous, albeit somewhat clumsy, yet universally accepted, expression for the partial derivative.

Here
$$du = ye^{xy}\,dx + xe^{xy}\,dy.$$
Now, this equation holds, no matter whether the independent variables be (x, y) or (r, s). In the latter case,*

$$dx = \frac{r}{r^2+s^2}\Delta r + \frac{s}{r^2+s^2}\Delta s, \qquad dy = \frac{-s}{r^2+s^2}\Delta r + \frac{r}{r^2+s^2}\Delta s.$$

Hence
$$du = \frac{ry - sx}{r^2+s^2}e^{xy}\Delta r + \frac{rx+sy}{r^2+s^2}e^{xy}\Delta s.$$

On the other hand,
$$du = \frac{\partial u}{\partial r}\Delta r + \frac{\partial u}{\partial s}\Delta s.$$

Thus these two right-hand sides are equal to each other, and Δr and Δs are independent variables. This can be true only when the coefficients of Δr are equal by themselves, and those of Δs are equal by themselves. For, we may set $\Delta s = 0$, $\Delta r \neq 0$, and then

$$\frac{\partial u}{\partial r}\Delta r = \frac{ry - sx}{r^2+s^2}e^{xy}\Delta r.$$

But $\Delta r \neq 0$, and so we can divide through by it. Thus we have:

$$\frac{\partial u}{\partial r} = \frac{ry - sx}{r^2+s^2}e^{xy}, \quad \text{and similarly,} \quad \frac{\partial u}{\partial s} = \frac{rx+sy}{r^2+s^2}e^{xy}.$$

In substance, this solution is the same as that of § 4; it differs only in form.

Example 2. If
$$u = f(x + a, y + b),$$
show that
$$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial a} \qquad \text{and} \qquad \frac{\partial u}{\partial y} = \frac{\partial u}{\partial b}.$$

Write
$$u = f(X, Y),$$
$$X = x + a, \qquad Y = y + b.$$

Then
$$du = f_1\,dX + f_2\,dY$$
$$= f_1\,dx + f_1\,da + f_2\,dy + f_2\,db.$$

But
$$du = \frac{\partial u}{\partial x}dx + \frac{\partial u}{\partial a}da + \frac{\partial u}{\partial y}dy + \frac{\partial u}{\partial b}db.$$

* It is immaterial whether we write Δr and Δs or dr and ds, since these infinitesimals are respectively equal to each other.

Here, $dx, \cdots, db$ are independent infinitesimals, and so the two expressions for du can be equal only when the corresponding coefficients of $dx, \cdots, db$ are respectively equal. Hence

$$\frac{\partial u}{\partial x}=f_1, \quad \frac{\partial u}{\partial a}=f_1, \quad \frac{\partial u}{\partial y}=f_2, \quad \frac{\partial u}{\partial b}=f_2,$$

and thus

$$\frac{\partial u}{\partial x}=\frac{\partial u}{\partial a}, \qquad \frac{\partial u}{\partial y}=\frac{\partial u}{\partial b}.$$

EXERCISES

Work Exercises 1–5 and 7, 8 of § 4 by the method of this paragraph.

7. Law of the Mean for Functions of Several Variables. Equation (4) of § 3 embodies a certain form of the law of the mean for functions of several variables; but there is a more symmetric form, which is easier to remember and is equally useful in practice.

Let $f(x, y)$ be continuous, together with its first partial derivatives, throughout the region

$$a \leqq x \leqq a+h, \qquad b \leqq y \leqq b+k.$$

Form the function

$$\Phi(t)=f(a+th, \quad b+tk), \qquad 0 \leqq t \leqq 1,$$

and apply to it the Law of the Mean, § 2:

$$\Phi(1)-\Phi(0)=\Phi'(\theta), \qquad 0<\theta<1.$$

Hence

(1) $\quad f(a+h,\, b+k)=f(a,\, b)+hf_1(a+\theta h,\, b+\theta k)+kf_2(a+\theta h,\, b+\theta k),$

where $\qquad 0<\theta<1.$

Equation (1) expresses the Law of the Mean for functions of several variables, which we set out to establish. h and k may, one or both, be negative; but θ always lies between 0 and 1. The extension to the case of a function of more than two variables is immediate.

8. Euler's Theorem for Homogeneous Functions. A function u is said to be *homogeneous* if, when each of the arguments is multiplied by one and the same quantity, the function is merely multiplied by a power of this quantity. For definiteness we will assume three arguments:

(1) $$u=f(x, y, z),$$

$$f(\lambda x, \lambda y, \lambda z)=\lambda^n f(x, y, z).$$

The explicit function shall be denoted by f:

(3) $$z = f(x, y).$$

Thus

(4) $$F[x, y, f(x, y)] \equiv 0,$$

i.e. this equation is an *identity*, since it holds for all values of the arguments, x and y.

The problem is to find $\partial z/\partial x$ or $f_1(x, y)$, and similarly $\partial z/\partial y$. To do this, let

(5) $$u = F(x, y, z).$$

Then

(6) $$du = F_1\, dx + F_2\, dy + F_3\, dz.$$

This equation holds, not only when x, y, z are the independent variables, but also when, in particular, x and y are the independent variables, z being replaced by the function (3); cf. § 5, Theorem. Under this hypothesis we have:

(*i*) $$u = \text{const.}\ (= 0).$$

For, $$u = F[x, y, f(x, y)],$$

and the value of the right-hand side, by (4), is 0. Hence

(7) $$du = 0.$$

(*ii*) $$dz = \frac{\partial z}{\partial x}\, dx + \frac{\partial z}{\partial y}\, dy,$$

where $\partial z/\partial x$ and $\partial z/\partial y$ are the derivatives we wish to find, and moreover dx and dy are independent infinitesimals.

Returning to (6) we now substitute for du its value from (7), and for dz its value from (*ii*); hence

$$0 = F_1 dx + F_2 dy + F_3\left(\frac{\partial z}{\partial x}\, dx + \frac{\partial z}{\partial y}\, dy\right),$$

or $$\left(F_1 + F_3 \frac{\partial z}{\partial x}\right) dx + \left(F_2 + F_3 \frac{\partial z}{\partial y}\right) dy = 0.$$

Since dx and dy are independent infinitesimals, it follows that each parenthesis, by itself, must vanish:

$$F_1 + F_3 \frac{\partial z}{\partial x} = 0, \qquad F_2 + F_3 \frac{\partial z}{\partial y} = 0.$$

Hence

(8) $$\frac{\partial z}{\partial x} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial z}},$$

with a similar equation for $\partial z/\partial y$.

This is the theorem concerning the differentiation of implicit functions, which we set out to prove. Observe that, in the differentiation on the left-hand side, the independent variables are x and y; whereas, in the differentiations on the right-hand side, the independent variables are x, y, and z.

Both theorem and proof apply to the case of any number of variables. Thus if

(9) $$F(x, y) = 0,$$

(10) $$\frac{dy}{dx} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial y}}.$$

It is often convenient to designate the dependent variable in the explicit form by u:

(11) $$\begin{cases} F(u, x, y, z, \cdots) = 0, \\ u = f(x, y, z, \cdots). \end{cases}$$

Then

(12) $$\frac{\partial u}{\partial x} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial u}},$$

with similar formulas for $\partial u/\partial y$, $\partial u/\partial z$, etc.

Example. Differentiate z partially, where

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1.$$

Here

$$F(x, y, z) = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1,$$

and we have:

$$\frac{2x}{a^2} + \frac{2z}{c^2}\frac{\partial z}{\partial x} = 0, \qquad \frac{\partial z}{\partial x} = -\frac{c^2 x}{a^2 z};$$

$$\frac{2y}{b^2} + \frac{2z}{c^2}\frac{\partial z}{\partial y} = 0, \qquad \frac{\partial z}{\partial y} = -\frac{c^2 y}{b^2 z}.$$

Remark. We might equally well have assumed the implicit equation in the form

$$F(u, x, y, z, \cdots) = C,$$

where C is a constant. The result, (12), would be of the same form.

Thus, in the above Example, we should have:

$$F(x, y, z) = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2},$$

the remainder of the work being as before.

EXERCISES

1. If $$pv^{1.41} = C, \qquad \text{find } \frac{dv}{dp}.$$

2. If z is defined by the equation

$$x^3 + y^3 = 3\,xyz,$$

find $\dfrac{\partial z}{\partial x}$.

3. If $$ye^{-xy} = \sin x, \qquad \text{find } \frac{dy}{dx}.$$

4. If $$u = f(xu, y),$$

show that $$\frac{\partial u}{\partial x} = \frac{uf_1(xu, y)}{1 - xf_1(xu, y)}.$$

5. If $$u = f(x - u, y - u), \qquad \text{find } \frac{\partial u}{\partial y}.$$

6. Show that, if $F(x, y, z) = 0$,

$$\frac{\partial^2 z}{\partial x^2} = -\frac{F_{33}F_1^2 - 2\,F_{13}F_1F_3 + F_{11}F_3^2}{F_3^3},$$

and compute $\dfrac{\partial^2 z}{\partial x\,\partial y}$ and $\dfrac{\partial^2 z}{\partial y^2}$.

10. Continuation. Simultaneous Equations. Let the functions u and v be defined by the simultaneous equations:

(1) $$F(u, v, x, y) = 0, \qquad \Phi(u, v, x, y) = 0,$$

where F and Φ are continuous, together with their partial derivatives of the first order, in the neighborhood of a point (u_0, v_0, x_0, y_0), and both vanish there:

(2) $$F(u_0, v_0, x_0, y_0) = 0, \qquad \Phi(u_0, v_0, x_0, y_0) = 0.$$

We will demand furthermore that the Jacobian determinant,

(3) $$J = \begin{vmatrix} \dfrac{\partial F}{\partial u} & \dfrac{\partial F}{\partial v} \\ \dfrac{\partial \Phi}{\partial u} & \dfrac{\partial \Phi}{\partial v} \end{vmatrix} = \frac{\partial (F, \Phi)}{\partial (u, v)},$$

be different from 0 in the above point. The explicit functions shall be denoted by f and ϕ:

(4) $$u = f(x, y), \qquad v = \phi(x, y).$$

Thus we have the identities:

(5) $$F[f(x, y), \phi(x, y), x, y] \equiv 0, \qquad \Phi[f(x, y), \phi(x, y), x, y] \equiv 0.$$

The problem is to compute the derivatives of the first order of the functions u and v, namely,

$$\frac{\partial u}{\partial x}, \quad \frac{\partial u}{\partial y}, \quad \frac{\partial v}{\partial x}, \quad \frac{\partial v}{\partial y}.$$

The procedure is similar to that of the preceding paragraph. Let

$$U = F(u, v, x, y), \qquad V = \Phi(u, v, x, y).$$

Then

(6) $$\begin{cases} dU = F_1\, du + F_2\, dv + F_3\, dx + F_4\, dy, \\ dV = \Phi_1\, du + \Phi_2\, dv + \Phi_3\, dx + \Phi_4\, dy. \end{cases}$$

These equations hold, no matter what the independent variables may be. In particular, then, we may replace u and v by the functions (4). We then have

(*i*) $$U = \text{const. } (= 0), \qquad V = \text{const. } (= 0),$$

and hence $$dU = 0, \qquad dV = 0;$$

(*ii*) $$du = \frac{\partial u}{\partial x} dx + \frac{\partial u}{\partial y} dy, \qquad dv = \frac{\partial v}{\partial x} dx + \frac{\partial v}{\partial y} dy.$$

On substituting these values in (6) and collecting terms, the following equations result:

(7) $$\left[F_1 \frac{\partial u}{\partial x} + F_2 \frac{\partial v}{\partial x} + F_3\right] dx + \left[F_1 \frac{\partial u}{\partial y} + F_2 \frac{\partial v}{\partial y} + F_4\right] dy = 0,$$

with a similar equation, in which F is replaced by Φ. Since dx and dy are independent infinitesimals, the individual brackets must all be 0. Hence

(8) $$\begin{cases} F_1 \dfrac{\partial u}{\partial x} + F_2 \dfrac{\partial v}{\partial x} + F_3 = 0, \\ \Phi_1 \dfrac{\partial u}{\partial x} + \Phi_2 \dfrac{\partial v}{\partial x} + \Phi_3 = 0, \end{cases}$$

with a similar pair of equations for $\partial u/\partial y$ and $\partial v/\partial y$.

On solving these equations we find:

$$(9) \qquad \frac{\partial u}{\partial x} = -\frac{\begin{vmatrix} F_3 & F_2 \\ \Phi_3 & \Phi_2 \end{vmatrix}}{\begin{vmatrix} F_1 & F_2 \\ \Phi_1 & \Phi_2 \end{vmatrix}}, \qquad \frac{\partial u}{\partial y} = -\frac{\begin{vmatrix} F_4 & F_2 \\ \Phi_4 & \Phi_2 \end{vmatrix}}{\begin{vmatrix} F_1 & F_2 \\ \Phi_1 & \Phi_2 \end{vmatrix}},$$

with similar equations for $\partial v/\partial x$ and $\partial v/\partial y$.

We have assumed two variables, x and y; but the number here is arbitrary. We may have a single variable, x; or we may have three variables, x, y, z; or any larger number. Moreover, both theorem and proof admit immediate extension to the case of n equations, defining a system of n functions. Thus, if

$$(10) \qquad \begin{cases} F(u, v, w, x, y, \cdots) = 0, \\ \Phi(u, v, w, x, y, \cdots) = 0, \\ \Psi(u, v, w, x, y, \cdots) = 0, \end{cases}$$

define the functions

$$u = f(x, y, \cdots), \qquad v = \phi(x, y, \cdots), \qquad w = \psi(x, y, \cdots),$$

we have:

$$(11) \qquad \frac{\partial u}{\partial x} = -\frac{\begin{vmatrix} F_4 & F_2 & F_3 \\ \Phi_4 & \Phi_2 & \Phi_3 \\ \Psi_4 & \Psi_2 & \Psi_3 \end{vmatrix}}{\begin{vmatrix} F_1 & F_2 & F_3 \\ \Phi_1 & \Phi_2 & \Phi_3 \\ \Psi_1 & \Psi_2 & \Psi_3 \end{vmatrix}}, \qquad \frac{\partial u}{\partial y} = -\frac{\begin{vmatrix} F_5 & F_2 & F_3 \\ \Phi_5 & \Phi_2 & \Phi_3 \\ \Psi_5 & \Psi_2 & \Psi_3 \end{vmatrix}}{\begin{vmatrix} F_1 & F_2 & F_3 \\ \Phi_1 & \Phi_2 & \Phi_3 \\ \Psi_1 & \Psi_2 & \Psi_3 \end{vmatrix}}, \qquad \text{etc.}$$

where the denominator is the Jacobian determinant with respect to u, v, w,

$$(12) \qquad J = \begin{vmatrix} \frac{\partial F}{\partial u} & \frac{\partial F}{\partial v} & \frac{\partial F}{\partial w} \\ \frac{\partial \Phi}{\partial u} & \frac{\partial \Phi}{\partial v} & \frac{\partial \Phi}{\partial w} \\ \frac{\partial \Psi}{\partial u} & \frac{\partial \Psi}{\partial v} & \frac{\partial \Psi}{\partial w} \end{vmatrix} = \frac{\partial (F, \Phi, \Psi)}{\partial (u, v, w)},$$

which is required to be different from 0.

Second Method. We have deduced equations (8) by the method of differentials. It is, however, possible to obtain them directly from the identities (5) by differentiating the latter equations partially with respect to x by the theorem of § 4. This method has the advantage that the computation is thus systematically arranged, and so there is less chance for numerical errors to creep in.

Example. Given the equations:

$$u^3 + xv = y,$$
$$v^3 + yu = x.$$

To find $\dfrac{\partial u}{\partial x}$.

Differentiating these equations with respect to x, we have:

$$3\,u^2\frac{\partial u}{\partial x} + x\frac{\partial v}{\partial x} + v = 0,$$

$$y\frac{\partial u}{\partial x} + 3\,v^2\frac{\partial v}{\partial x} = 1.$$

Hence
$$\frac{\partial u}{\partial x} = \frac{x + 3\,v^3}{xy - 9\,u^2v^2}.$$

EXERCISES

1. Compute $\partial u/\partial y$ in the Example of the text.

2. If $\left\{\begin{array}{l} xu^2 + yv^3 = xy^2, \\ yu^3 - xv^2 = x^2y, \end{array}\right.$ find $\dfrac{\partial v}{\partial y}$.

3. If $\left\{\begin{array}{l} xu + uv - v = x^3, \\ v^2 - xv + u = x^2, \end{array}\right.$ find $\dfrac{du}{dx}$.

4. Compute $\partial v/\partial x$ and $\partial v/\partial y$ in the Example of the text from Equations (9).

5. Compute $\partial v/\partial x$, $\partial v/\partial y$, $\partial v/\partial z$ from equations (10).

11. The Inverse of a Transformation. The idea of a *transformation* of a plane has been set forth in the *Analytic Geometry*, Chap. XV, p. 330, and the student should be familiar with the examples there discussed and with their analytic treatment. Moreover, he should of his own initiative extend the simpler of these examples to space of three dimensions, availing himself of the treatment of the analytically kindred problem of the transformation of coordinates, given in the last chapter, p. 592.

These examples illustrate what is meant in the general case by the *transformation* of a region S of one plane on a region Σ of the same or a second plane. Let $P: (u, v)$ be any point of S, and let

$Q:(x, y)$ be the point of Σ into which P is carried by the given transformation. Then x and y are *functions* of u and v:

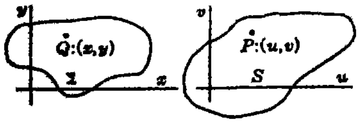


FIG. 35

$$(1) \qquad \begin{cases} x = f(u, v), \\ y = \phi(u, v). \end{cases}$$

We will assume these functions to be continuous, together with their first partial derivatives, and we will demand, furthermore, that the Jacobian determinant,

$$J = \begin{vmatrix} \dfrac{\partial f}{\partial u} & \dfrac{\partial f}{\partial v} \\ \dfrac{\partial \phi}{\partial u} & \dfrac{\partial \phi}{\partial v} \end{vmatrix} = \frac{\partial(x, y)}{\partial(u, v)},$$

vanish at no point of S.

Since by hypothesis the relation of the points of S to those of Σ is one-to-one, an arbitrary point $Q:(x, y)$ of Σ leads to one and only one point, P, of S. Hence the coordinates of P, namely, u and v, are functions of those of Q, *i.e.* x and y:

$$(2) \qquad \begin{cases} u = F(x, y), \\ v = \Phi(x, y), \end{cases}$$

and these functions are continuous, together with their first partial derivatives, as will be shown in detail later; § 12.

The pair of equations (2) expresses explicitly the transformation of the region Σ on the region S. This transformation is called the *inverse* of the given transformation, which is expressed analytically by equations (1).

Even when f and ϕ are simple functions of u and v, it is often impossible to express F and Φ in terms of the functions with which we are familiar, and so the derivatives of F and Φ must be computed indirectly. This can be done by the method of § 10, for we need merely set

$$(3) \qquad f(u, v) - x = 0, \qquad \phi(u, v) - y = 0.$$

It is instructive, however, to apply the second method there set forth. If we substitute for u and v in the first equation (1) the functions $F(x, y)$ and $\Phi(x, y)$ respectively, we have an identity in x and y:

$$(4) \qquad x \equiv f[F(x, y), \Phi(x, y)],$$

i.e. in the last analysis, the equation $x = x$. We may, therefore, differentiate each side of this equation, which we will retain in the form (1), partially with respect to x, according to the theorem of § 4:

(5) $$1=\frac{\partial x}{\partial u}\frac{\partial u}{\partial x}+\frac{\partial x}{\partial v}\frac{\partial v}{\partial x},$$

and it is important to understand what these partial derivatives mean. In $\partial x/\partial u$, the independent variables are u and v, and the dependent variable is the function $x=f(u, v)$. In $\partial u/\partial x$, the independent variables are x and y, and the dependent variable is the function $u=F(x, y)$. Similarly for the second term.

Equation (5) is an equation for two of the unknown derivatives, $\partial u/\partial x$ and $\partial v/\partial x$, which we are trying to find, in terms of derivatives of the given functions. A second equation connecting these unknown derivatives can be found by the aid of the second equation (1). Here we have the identity in x and y:

(6) $$y\equiv\phi[F(x,y),\ \Phi(x,y)].$$

Hence, differentiating this equation, written in the form (1), partially with respect to x, we have:

(7) $$0=\frac{\partial y}{\partial u}\frac{\partial u}{\partial x}+\frac{\partial y}{\partial v}\frac{\partial v}{\partial x}.$$

Equations (5) and (7) are a pair of simultaneous linear equations for the two unknowns, $\partial u/\partial x$ and $\partial v/\partial x$, and they can be solved by the methods of high school algebra.

Example. Let the given transformation be represented by the equations:

(*i*) $$\begin{cases} x=3u-8v, \\ y=2u-5v, \end{cases}$$

corresponding to (1). The inverse transformation will then be given by the equations:

(*ii*) $$\begin{cases} u=-5x+8y, \\ v=-2x+3y, \end{cases}$$

corresponding to (2).

To find the derivatives by the method set forth in the text, differentiate each of the equations (*i*) partially with respect to x:

$$\begin{cases} 1=3\dfrac{\partial u}{\partial x}-8\dfrac{\partial v}{\partial x} \\ 0=2\dfrac{\partial u}{\partial x}-5\dfrac{\partial v}{\partial x} \end{cases}$$

On solving these equations for $\partial u/\partial x$ and $\partial v/\partial x$ we find:

$$\frac{\partial u}{\partial x} = -5, \qquad \frac{\partial v}{\partial x} = -2.$$

The result agrees with that which can here be obtained directly by differentiating the equations (*ii*) with respect to x.

The student should verify equations (4) and (6) in the case of this example.

EXERCISES

1. If
$$\begin{cases} x = u + v, \\ y = u^2 + v^3, \end{cases}$$
find $\dfrac{\partial v}{\partial x}$.

2. If
$$\begin{cases} x = u + v + w, \\ y = u^2 + v^2 + w^2, \\ z = u^3 + v^3 + w^3, \end{cases}$$
show that
$$\frac{\partial u}{\partial x} = \frac{vw(w - v)}{(u - v)(v - w)(w - u)},$$
and compute $\dfrac{\partial u}{\partial y}$, $\dfrac{\partial u}{\partial z}$.

3. If
$$\begin{cases} x = u + ve^u, \\ y = u, \end{cases}$$
find $\partial v/\partial y$ by the method of the text.

Solve the given equations for u and v, and verify the result obtained for $\partial v/\partial y$ by direct computation of this derivative.

4. Verify equations (4) and (6) of the text for the case of the transformation given in the preceding question.

12. An Existence Theorem. In § 9 we have considered an implicit function defined by the equation

$$F(u, x, y) = 0, \tag{1}$$

and we have assumed that this function,

$$u = f(x, y), \tag{2}$$

is continuous together with its first partial derivatives. Very simple examples suffice to show that equation (1) may not give rise to such a function as (2). *Thus the equation*

$$F(u, x, y) = u^2 + x^2 + y^2 + 1 = 0 \tag{3}$$

has no roots, for the sum of three squares cannot be equal to -1. Again, the equation

$$(4)\qquad F(u, x, y) = u^2 + x^2 + y^2 = 0$$

admits only the single solution, (0, 0, 0), and so, again, fails to define a function (2). Finally, the equation

$$(5)\qquad F(u, x, y) = (u - x^2)^2 + y^2 = 0$$

defines, not a *surface* as represented by (2), but a *curve*, namely,

$$(6)\qquad u = x^2, \qquad y = 0.$$

Is it possible to tell from simple properties of the given function $F(u, x, y)$ when equation (1) will define a function (2) with the properties presupposed in § 9? The answer is affirmative, and is given by the following theorem.

THEOREM. *Let $F(u, x, y)$ be continuous, together with its first partial derivatives, throughout a certain neighborhood of a given point (u_0, x_0, y_0). Let F vanish at this point, but let $\partial F/\partial u$ be different from 0:*

$$(7)\qquad F(u_0, x_0, y_0) = 0, \qquad F_u(u_0, x_0, y_0) \neq 0.$$

Then there is a function of x and y,

$$(8)\qquad u = f(x, y),$$

continuous throughout a certain neighborhood of the point (x_0, y_0) and taking on the value u_0 there:

$$u_0 = f(x_0, y_0),$$

which function satisfies the equation

$$(9)\qquad F(u, x, y) = 0;$$

i.e. if $f(x, y)$ be substituted for u in the given function $F(u, x, y)$, the latter vanishes identically:

$$(10)\qquad F[f(x, y), x, y] \equiv 0.$$

Moreover, the only triples of values, (u, x, y), which lie in the neighborhood of the point (u_0, x_0, y_0) and satisfy equation (9) are those which are connected by the equation (8).

Finally, this function, $f(x, y)$, has continuous first derivatives.

This theorem would seem in one respect to be unsatisfactory, in that it does not tell us how large these "neighborhoods" are. The neighborhood of the point (u_0, x_0, y_0) may be taken as the points of a certain rectangular parallelepiped:

$$|x - x_0| < A, \qquad |y - y_0| < A, \qquad |u - u_0| < B,$$

where A and B are two constants which may have to be chosen very small. And then the function $f(x, y)$ is to be considered only for the points of a certain square,

$$|x - x_0| < h, \qquad |y - y_0| < h,$$

where h is surely not greater than A, and may be less. These restrictions, on the one hand, lie in the nature of the case. The theorem is not true in general if they be removed. On the other hand, the theorem, restricted as it is, is nevertheless exceedingly useful in practice.

Of course, the number of variables $(x, y, \cdots)$ is arbitrary. If F depends only on two, $F(u, x)$, then f will depend on one,

$$u = f(x).$$

And if F depends on $n + 1$, $F(u, x_1, \cdots, x_n)$, f will depend on n:

$$u = f(x_1, \cdots, x_n).$$

A proof of this theorem and the following will be found in Goursat-Hedrick, *Mathematical Analysis*, Vol. I, Chap. II, § 20, and in the author's *Funktionentheorie*, Vol. I, Chap. II, § 4. The latter treatment lays stress on the geometrical interpretation of the analysis employed.

The partial derivatives of u are computed by the method set forth above, in § 9, and the formulas (8), (10), (12) there obtained give the solution of this part of the problem.

Geometric Evidence. In the simplest case, namely, that in which the number of variables is two, and it is thus a question of showing that the equation

$$F(u, x) = 0 \tag{11}$$

determines a curve,

$$u = f(x), \tag{12}$$

it is possible to make the truth of the theorem plausible geometrically as follows. Consider the surface,

$$z = F(u, x).$$

This surface meets the coordinate plane $z = 0$ in the point $(u, x, z) = (u_0, x_0, 0)$, and it is not tangent to it there, but actually cuts it at an angle. For, the direction components of the normal to the surface are proportional to $F_u(u, x)$, $F_x(u, x)$, -1, and since $F_u(u_0, x_0) \neq 0$, the normal cannot be perpendicular to the axis of u. Since the surface, moreover, is smooth in the neighborhood of the point in question, we should expect any oblique plane to cut it in a smooth curve. Finally, we see, if this be granted, that the curve is not perpendicular to the x-axis at the point in question, and so it is reasonable that its equation be expressible in the form (12).

The case of the inverse of a given function, which we have met repeatedly in the study of the Calculus, could be related directly to the foregoing, but is better dealt with as follows. Let x be given as a function of u,

(13) $$x = \phi(u),$$

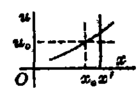


FIG. 36

where $\phi(u)$ and its derivative, $\phi'(u)$, are continuous at $u = u_0$, and $\phi'(u_0) \neq 0$. Then the equation defines u as a function of x, continuous and having a continuous first derivative in the neighborhood of the point $x = x_0$, where $x_0 = \phi(u_0)$. For, the curve which is the graph of equation (13), u being plotted as the ordinate and x as the abscissa, has a tangent at the point (u_0, x_0) which is not parallel to the u-axis. Hence a line $x = x'$, where x' differs but slightly from x_0, will cut this curve in just one point, and the ordinate of this point will be the inverse function

(14) $$u = f(x).$$

Simultaneous Equations. The theorem admits extension to a simultaneous system of p equations which determine p implicit functions. For definiteness, we state it for the particular case, $p = 3$.

THEOREM. *Let the functions*

$$F(u, v, w, x, y, \cdots), \qquad \Phi(u, v, w, x, y, \cdots), \qquad \Psi(u, v, w, x, y, \cdots)$$

be continuous, together with their first partial derivatives, throughout a region,

(13) $$|u - u_0| < B, \quad |v - v_0| < B, \quad |w - w_0| < B,$$
$$|x - x_0| < A, \quad |y - y_0| < A, \quad \cdots,$$

and let them all vanish at the point $(u_0, v_0, w_0, x_0, y_0, \cdots)$; *let the Jacobian determinant,*

$$J = \frac{\partial(F, \Phi, \Psi)}{\partial(u, v, w)},$$

be different from 0 *there. Then there exist three functions*

(14) $u = f(x, y, \cdots)$, $v = \phi(x, y, \cdots)$, $w = \psi(x, y, \cdots)$,

each continuous throughout a region

(15) $|x - x_0| < h$, $|y - y_0| < h$, $\cdots$, $h \leq A$,

and taking on the respective values u_0, v_0, w_0 *in the point* $(x_0, y_0, \cdots)$, *which functions satisfy the simultaneous equations*

(16) $F(u, v, w, x, y, \cdots) = 0$, $\Phi(u, v, w, x, y, \cdots) = 0$,

$$\Psi(u, v, w, x, y, \cdots) = 0.$$

Moreover, the only sets of values $(u, v, w, x, y \cdots)$ *which lie in the region* (13) *and satisfy the equations* (16) *simultaneously are (provided A and B are suitably restricted) such as are given by equations* (14).

Finally, the functions (14) *have continuous first partial derivatives.*

The derivatives of the functions (14) can be computed by the method set forth in § 10.

The Inverse of a Transformation. Let the transformation

(17) $x = f(u, v)$, $y = \phi(u, v)$

be given, where f and ϕ are continuous, together with their first partial derivatives, throughout the neighborhood of a point (u_0, v_0), and let x_0, y_0 denote respectively the values of these functions at this point. Let the Jacobian determinant,

$$J = \frac{\partial(f, \phi)}{\partial(u, v)},$$

be different from 0 at (u_0, v_0). Then the inverse of the transformation (17) is represented, in the neighborhood of the point (u_0, v_0) and the point (x_0, y_0), by two equations,

(18) $u = F(x, y)$, $v = \Phi(x, y)$,

where $F(x, y)$ and $\Phi(x, y)$, together with their first partial derivatives, are continuous throughout a certain neighborhood of the point (x_0, y_0).

The theorem holds for a similar transformation in any number of variables:

(19) $x_i = f_i(u_1, \cdots, u_n)$, $i = 1, 2, \cdots, n.$

13. Concerning Jacobians. The Jacobian of a set of p functions of p independent variables, as $(p=3)$:

(1) $$u=f(x, y, z), \qquad v=\phi(x, y, z), \qquad w=\psi(x, y, z),$$

namely, the determinant

$$J=\begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} & \frac{\partial w}{\partial z} \end{vmatrix},$$

is often represented, as we have repeatedly had occasion to remark, by the notation

(2) $$J=\frac{\partial(u, v, w)}{\partial(x, y, z)}.$$

If $$u_i=f_i(y_1, \cdots, y_n), \qquad i=1, \cdots, n,$$

and $$y_i=\phi_i(x_1, \cdots, x_n), \qquad i=1, \cdots, n,$$

it can be shown that

(3) $$\frac{\partial(u_1, \cdots, u_n)}{\partial(x_1, \cdots, x_n)}=\frac{\partial(u_1, \cdots, u_n)}{\partial(y_1, \cdots, y_n)}\frac{\partial(y_1, \cdots, y_n)}{\partial(x_1, \cdots, x_n)}.$$

Cf. Jordan, *Cours d'analyse*, Vol. I, 3d edition, 1893, p. 89. It is assumed that the functions f_i and ϕ_i are continuous, together with their first partial derivatives.

If the equations

(4) $$x_i=f_i(y_1, \cdots, y_n), \qquad i=1, \cdots, n,$$

represent a transformation with non-vanishing Jacobian

$$j=\frac{\partial(x_1, \cdots, x_n)}{\partial(y_1, \cdots, y_n)},$$

and if the inverse transformation be given by the equations

(5) $$y_i=\phi_i(x_1, \cdots, x_n), \qquad i=1, \cdots, n,$$

then the Jacobian of this transformation,

$$J=\frac{\partial(y_1, \cdots, y_n)}{\partial(x_1, \cdots, x_n)},$$

is the reciprocal of j:

(6) $$jJ=1, \qquad \text{or} \qquad J=\frac{1}{j}.$$

For here $$u_i=x_i, \qquad i=1, \cdots, n,$$

and thus the left-hand side of (3) reduces to unity.

Geometric Meaning of the Vanishing of the Jacobian. Consider the equations

$$u = f(x, y), \qquad v = \phi(x, y),$$

where f and ϕ are continuous, together with their first partial derivatives, throughout the neighborhood of the point (x_0, y_0), and let $u_0 = f(x_0, y_0)$, $v_0 = \phi(x_0, y_0)$. If the partial derivatives of f do not both vanish there, the equation

$$f(x, y) - u = 0,$$

where u is a parameter to which are assigned values near u_0, yields a one-parameter family of curves coursing the neighborhood in question, one and only one curve going through any given point of this region, and each curve being smooth and free from multiple points. For, the above equation can be solved for y in terms of x and u, or else for x in series of y and u, the function thus obtained being continuous.

Similarly, if the partial derivatives of ϕ are not both 0, the equation

$$\phi(x, y) - v = 0$$

represents a second family of the same character.

The curves

$$f(x, y) - u_0 = 0, \qquad \phi(x, y) - v_0 = 0$$

will be tangent to each other at (x_0, y_0) if and only if

$$f_x : \phi_x = f_y : \phi_y$$

there, *i.e.* if and only if the Jacobian

$$J = \begin{vmatrix} \dfrac{\partial f}{\partial x} & \dfrac{\partial f}{\partial y} \\ \dfrac{\partial \phi}{\partial x} & \dfrac{\partial \phi}{\partial y} \end{vmatrix}$$

vanishes there.

The assumption that $J \neq 0$ at (x_0, y_0) carries with it, because of the continuity of this function, that J does not vanish at any other point of a suitably restricted neighborhood of (x_0, y_0), and hence the two curves, one from each family, which go through an arbitrary point of this region cut each other at an angle which is different from 0 or π. It is plausible geometrically, therefore, that these curves can have no second point of intersection in the neighborhood of (x_0, y_0); for, the directions of the curves of one family vary only slightly from one another; and similarly for the other family. But this means precisely that the equations (1) admit a single-valued inverse. For it says that, to an arbitrary pair of values (u, v) near (u_0, v_0), there corresponds one and only one pair of values (x, y) near (x_0, y_0).

If three equations be given:

$$u = f(x, y, z), \qquad v = \phi(x, y, z), \qquad w = \psi(x, y, z),$$

where f, ϕ, ψ are continuous, together with their first partial derivatives, in the neighborhood of (x_0, y_0, z_0) and the derivatives of no one of the functions are all zero at this point, then each of the equations represents a one-parameter family of surfaces coursing the region in question, one and only one surface going through each point of the region.

The vanishing of the Jacobian,

$$J = \frac{\partial(u, v, w)}{\partial(x, y, z)},$$

here signifies that all three surfaces are tangent to the same line. For, if no two of the surfaces are tangent to each other at the point in question, the direction components of the curve of intersection of the first two will be:

$$\begin{vmatrix} f_2 & f_3 \\ \phi_2 & \phi_3 \end{vmatrix}, \qquad \begin{vmatrix} f_3 & f_1 \\ \phi_3 & \phi_1 \end{vmatrix}, \qquad \begin{vmatrix} f_1 & f_2 \\ \phi_1 & \phi_2 \end{vmatrix};$$

and if $J = 0$, then

$$\psi_1 \begin{vmatrix} f_2 & f_3 \\ \phi_2 & \phi_3 \end{vmatrix} + \psi_2 \begin{vmatrix} f_3 & f_1 \\ \phi_3 & \phi_1 \end{vmatrix} + \psi_3 \begin{vmatrix} f_1 & f_2 \\ \phi_1 & \phi_2 \end{vmatrix} = 0,$$

or, the normal to the third surface is perpendicular to this curve.

We see, then, that, if $J \neq 0$, no two of the surfaces can be tangent, and the curve of intersection of two of the surfaces cuts the third surface obliquely. And now reasoning similar to that of the foregoing case leads to the inference that an arbitrary set of values (u, v, w) near (u_0, v_0, w_0) gives rise to surfaces which cut in one and only one point in the neighborhood of (x_0, y_0, z_0). Hence the inverse transformation is also single-valued.

In the case of a transformation (4) with $n > 3$, the geometric evidence is lacking, since we should need a space of $n > 3$ dimensions. The analytic proof, however, applies equally well, no matter how large n may be; cf. Goursat-Hedrick, *Mathematical Analysis*, Vol. I, p. 45, § 25, or the author's *Funktionentheorie*, Vol. I, Chap. II, §§ 5–7.

Identical Vanishing of the Jacobian. If the Jacobian of the functions (1) vanishes identically, these functions are connected by a relation

$$\Omega(u, v, w) = 0,$$

where Ω is continuous, together with its first partial derivatives. The theorem holds for any number of variables. Cf. Jordan, *l. c.*, and the author's *Funktionentheorie*, Vol. II, p. 122.

14. A Question of Notation. *Problem.* Suppose

$$u = f(x, y), \qquad y = \phi(x, z);$$

to find $\dfrac{\partial u}{\partial x}$.

Before beginning a partial differentiation the first question which we must ask ourselves is: *What are the independent variables?* Hitherto the notation has always been such as to suggest readily what the independent variables are. In the present case they may be:

(*a*) x and y; or (*b*) x and z; or (*c*) y and z.

We can indicate which case is meant by writing the independent variables as subscripts, thus:

$$(a) \quad \frac{\partial u_{xy}}{\partial x}; \qquad (b) \quad \frac{\partial u_{xz}}{\partial x}.$$

In case (*c*) $\dfrac{\partial u}{\partial x}$ has no meaning.

In case (*a*), $$\frac{\partial u_{xy}}{\partial x} = f_1(x, y).$$

Case (*b*) can be brought directly under the method of § 4 by introducing a new letter in representing the independent variables of the second class:

$$u = f(x, y), \qquad \begin{cases} \text{independent variables of} \\ \text{the 1st class, } (x, y); \end{cases}$$

$$x = r, \qquad y = \phi(r, z), \qquad \begin{cases} \text{independent variables of} \\ \text{the 2d class, } (r, z). \end{cases}$$

Thus $\dfrac{\partial u_{xz}}{\partial x}$ now becomes $\dfrac{\partial u}{\partial r}$ and can be computed in the usual manner:

$$\frac{\partial u}{\partial r} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial r},$$

or

$$\frac{\partial u_{xz}}{\partial x} = \frac{\partial u_{xy}}{\partial x} + \frac{\partial u_{xy}}{\partial y}\frac{\partial \phi}{\partial x}.$$

In Thermodynamics, the pressure, p, the volume, v, and the temperature, t, are connected by the so-called *characteristic equation*,

$$\phi(p, v, t) = 0,$$

and so only two of these three variables can be chosen as independent. It may happen that the pressure can be expressed as a function of the temperature and the energy, E:

$$p = F(t, E),$$

and that it is convenient to express the volume by the characteristic equation as a function of t and p:

$$v = \Phi(t, p).$$

A notation which the physicists use to express Φ_1 and Φ_2 is:

$$\left(\frac{\partial v}{\partial t}\right)_p \quad \text{for} \quad \frac{\partial v_{tp}}{\partial t} = \Phi_1(t, p);$$

$$\left(\frac{\partial v}{\partial p}\right)_t \quad \text{for} \quad \frac{\partial v_{tp}}{\partial p} = \Phi_2(t, p),$$

the subscript indicating the variable which is held fast.

EXERCISES

1. If
$$u = 2xy$$
and
$$2x + 3y + 5z = 1,$$
explain all the meanings which $\frac{\partial u}{\partial x}$ may have, and evaluate this derivative in each case.

2. Show that
$$\left(\frac{\partial v}{\partial p}\right)_t = -\frac{\phi_1(p, v, t)}{\phi_2(p, v, t)}.$$

3. Prove:
$$\left(\frac{\partial v}{\partial t}\right)_E = \left(\frac{\partial v}{\partial t}\right)_p + \left(\frac{\partial v}{\partial p}\right)_t \left(\frac{\partial p}{\partial t}\right)_E.$$

15. Small Errors. In the case of functions of a single variable we have seen that the linear term in the expansion of Taylor's Theorem:

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \cdots,$$

can frequently be used to express with sufficient accuracy the effect of a small error of observation on the final result, cf. *Introduction to* *the Calculus*, p. 417, § 16. This term, $f'(x_0)(x-x_0)$, is precisely the differential of the function, df, for $x=x_0$.

The differential of a function of several variables can be used for a similar purpose. If $x, y, \cdots$ are the observed quantities and u the magnitude to be computed, then the precise error in u due to errors of observation $\Delta x = dx$, $\Delta y = dy$, etc. is Δu. But

$$du = \frac{\partial u}{\partial x}dx + \frac{\partial u}{\partial y}dy + \cdots$$

will frequently differ from Δu by a quantity so small that either is as accurate as the observations will warrant, — and du is more easily computed.

Example. The period of a simple pendulum is

$$T = 2\pi\sqrt{\frac{l}{g}}.$$

To find the error caused by errors in measuring l and g, or in the variation of l due to temperature and of g due to the location on the earth's surface.

Here
$$dT = \frac{\pi}{\sqrt{lg}}dl - \frac{\pi}{g}\sqrt{\frac{l}{g}}dg,$$
or
$$\frac{dT}{T} = \frac{1}{2}\frac{dl}{l} - \frac{1}{2}\frac{dg}{g},$$

and hence a small positive error of k per cent in observing l will increase the computed time by $\frac{1}{2}k$ per cent, and a small positive error of k' per cent in the value of g will decrease the computed time by $\frac{1}{2}k'$ per cent.

EXERCISES

1. A side c of a triangle is determined in terms of the other two sides and the included angle by means of the formula:

$$c^2 = a^2 + b^2 - 2ab\cos\omega.$$

Find approximately the error in c due to slight errors in measuring a, b, and ω. *Ans.* The percentage error is given by the formula:

$$\frac{dc}{c} = \frac{(a - b\cos\omega)\,da + (b - a\cos\omega)\,db + ab\sin\omega\,d\omega}{a^2 + b^2 - 2ab\cos\omega}.$$

2. Find approximately the error in the computed area of the triangle in the preceding question.

3. The acceleration of gravity as determined by an Atwood's machine is given by the formula:

$$g = \frac{2s}{t^2}.$$

Find approximately the error due to small errors in observing s and t.

4. Describe an experiment you have performed to determine the focal length of a lens, recall the relative degrees of accuracy you attained in the successive observations, and discuss the effects of the errors of observation on the final result.

16. Directional Derivatives. Let a function

$$u = f(x, y)$$

be given at each point of a region S of the x, y plane and let a curve C be given passing through a point $P: (x_0, y_0)$ of the region. Let P' be a second point of C, and form the quotient:

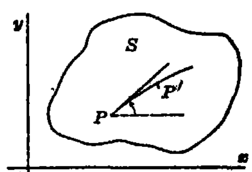


FIG. 37

$$\frac{u_{P'} - u_P}{\overset{\frown}{PP'}}.$$

The limit of this quotient, when P' approaches P, is defined as the *directional derivative* of u along the curve C. We set $u_{P'} - u_P = \Delta u$, $\overset{\frown}{PP'} = \Delta\xi$, and write

$$\lim_{\Delta\xi \doteq 0} \frac{\Delta u}{\Delta \xi} = \frac{\partial u}{\partial \xi}.$$

If, in particular, C is a ray parallel to the axis of x and having the same sense, the directional derivative has the value of the partial derivative, $\partial u/\partial x$; if the ray has the opposite sense, the directional derivative is equal to $-\partial u/\partial x$. A similar remark applies to the axis of y.

To compute the directional derivative in the general case we make use of the Lemma of § 3; hence

$$\lim_{\Delta\xi \doteq 0} \frac{\Delta u}{\Delta \xi} = \frac{\partial u}{\partial x}\left(\lim_{\Delta\xi \doteq 0} \frac{\Delta x}{\Delta \xi}\right) + \frac{\partial u}{\partial y}\left(\lim_{\Delta\xi \doteq 0} \frac{\Delta y}{\Delta \xi}\right),$$

or

(1) $$\frac{\partial u}{\partial \xi} = \frac{\partial u}{\partial x}\cos\alpha + \frac{\partial u}{\partial y}\sin\alpha.$$

The extension of the definition to space of three dimensions is immediate. We have:

$$\frac{\partial u}{\partial \xi} = \frac{\partial u}{\partial x}\cos\alpha + \frac{\partial u}{\partial y}\cos\beta + \frac{\partial u}{\partial z}\cos\gamma, \tag{2}$$

where α, β, γ are the direction angles of C at P.

EXERCISES

1. If a normal be drawn to a plane curve at any point P and if r denote the distance of a variable point of the plane from a fixed point O; γ, the angle between PO and the direction of the normal, show that

$$\frac{\partial r}{\partial n} = -\cos\gamma. \tag{3}$$

2. Explain the meaning of $\dfrac{\partial n}{\partial r}$ and show that

$$\frac{\partial n}{\partial r} = \frac{\partial r}{\partial n}. \tag{4}$$

17. Potential. Let a particle of mass m be situated at the fixed point $A:(a, b, c)$, and let a second particle, of mass unity, be situated at the variable point $P:(x, y, z)$, distant r from A. Then the force with which these particles attract each other, measured in gravitational units, will be

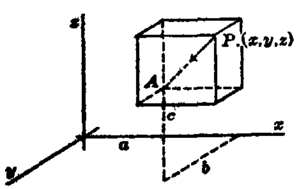


FIG. 38

$$F = \frac{m}{r^2}. \tag{1}$$

Consider, in particular, the force which the particle at A exerts on the particle at P. Its components along the axes will be

$$X = -m\frac{x-a}{r^3}, \qquad Y = -m\frac{y-b}{r^3}, \qquad Z = -m\frac{z-c}{r^3}. \tag{2}$$

If we consider an arbitrary direction from P, whose direction angles are α, β, γ, the component of the attraction along this direction will be

$$\Xi = F\cos\epsilon,$$

where ϵ denotes the angle between the direction and PA. Since by the cosine formula (*Analytic Geometry*, p. 426)

$$\cos\epsilon = -\frac{x-a}{r}\cos\alpha - \frac{y-b}{r}\cos\beta - \frac{z-c}{r}\cos\gamma,$$

we have

$$\Xi = X\cos\alpha + Y\cos\beta + Z\cos\gamma. \tag{3}$$

Let us form the function

$$(4) \qquad u = \frac{m}{r},$$

known as the *potential* of the mass m. Its partial derivatives are seen to give precisely the components of the attraction along the axes:

$$(5) \qquad \frac{\partial u}{\partial x} = -\frac{m}{r^2}\frac{\partial r}{\partial x},$$

$$r^2 = (x-a)^2 + (y-b)^2 + (z-c)^2,$$

$$2r\frac{\partial r}{\partial x} = 2(x-a), \qquad \frac{\partial r}{\partial x} = \frac{x-a}{r};$$

hence $$\frac{\partial u}{\partial x} = \left(-\frac{m}{r^2}\right)\left(\frac{x-a}{r}\right) = -m\frac{x-a}{r^3} = X.$$

Similarly,

$$\frac{\partial u}{\partial y} = Y, \qquad \frac{\partial u}{\partial z} = Z.$$

On substituting these values of X, Y, Z in (3), the right-hand side becomes

$$\frac{\partial u}{\partial x}\cos\alpha + \frac{\partial u}{\partial y}\cos\beta + \frac{\partial u}{\partial z}\cos\gamma.$$

But this is, by (2), § 16, precisely the directional derivative along the given direction, or $\frac{\partial u}{\partial \xi}$. Hence

$$(6) \qquad \Xi = \frac{\partial u}{\partial \xi},$$

and we thus have the theorem: *The component of the attraction due to the mass m, situated at A, is given by the directional derivative of the potential function* (4), *taken along the direction in question.*

The Case of n Masses. If, instead of a single mass at A, there are n masses, $m_1, \cdots, m_n$, situated respectively at the points $A_1, \cdots, A_n$, then the *potential* of these masses at the point P is defined as the function u, where

$$(7) \qquad u = \frac{m_1}{r_1} + \frac{m_2}{r_2} + \cdots + \frac{m_n}{r_n} = \sum_{k=1}^{n} \frac{m_k}{r_k},$$

r_k denoting the distance from P to $A_k : (a_k, b_k, c_k)$. Thus

$$r_k^2 = (x-a_k)^2 + (y-b_k)^2 + (z-c_k)^2.$$

The components of the attraction on a unit particle at P are given by the same formulas as before, namely:

$$\frac{\partial u}{\partial x} = X, \qquad \frac{\partial u}{\partial y} = Y, \qquad \frac{\partial u}{\partial z} = Z; \tag{8}$$

$$\Xi = \frac{\partial u}{\partial \xi}. \tag{10}$$

It will be shown in Chap. XI, § 2 that the change in the value of the potential function, when the unit particle at P describes an arbitrary path, is equal to the work done on this particle by the forces exerted by the n masses, and hence this function u is sometimes called the *force function*.

EXERCISE

Show that the potential function satisfies Laplace's Equation:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0.$$

The student will find elaborate applications of partial differentiation, taken from life (differential geometry, mathematical physics), in Goursat-Hedrick's *Mathematical Analysis*, Vol. I, Chap. II.

EXERCISES ON CHAPTER V

1. If

$$u = \frac{\cos y}{x}$$

and

$$x = r^2 - s, \qquad y = e^s,$$

find $\dfrac{\partial u}{\partial s}$.

2. If

$$u = e^{x \sin y} + x \log(x + y)$$

and

$$x = pqr, \qquad y = r \sin^{-1}(qr),$$

find $\dfrac{\partial u}{\partial q}$.

3. If

$$yx^y = \sin x,$$

find $\dfrac{dy}{dx}$.

4. If

$$\begin{cases} u^5 + v^5 + x^5 = 3y, \\ u^3 + v^3 + y^3 = -3x, \end{cases}$$

find $\dfrac{\partial u}{\partial x}$.

5. If
$$V = 2uv$$
and
$$\begin{cases} u^5 + v^5 + x^5 = 3y, \\ u^3 + v^3 + y^3 = -3x, \end{cases}$$
find $\dfrac{\partial V}{\partial x}$.

6. If
$$\begin{cases} u^3 + xv = y, \\ v^3 + yu = x, \end{cases}$$
find $\dfrac{\partial v}{\partial y}$.

7. If
$$\begin{cases} ue^v + vx = y \sin u, \\ u \cos u = x^2 + y^2, \end{cases}$$
find $\dfrac{\partial v}{\partial y}$.

8. If
$$\begin{cases} xu + uv = v + x, \\ v^2 - xv = u - x^2, \end{cases}$$
find $\dfrac{du}{dx}$.

9. If
$$\begin{cases} x = u + v + w, \\ y = u^2 + v^2 + w^2, \\ z = u^3 + v^3 + w^3, \end{cases}$$
show that
$$\frac{\partial u}{\partial x} = \frac{vw(w - v)}{(u - v)(v - w)(w - u)},$$
and find $\dfrac{\partial u}{\partial y}$

10. If
$$\begin{cases} x = u + v + w, \\ y = uv + vw + wu, \\ z = uvw, \end{cases}$$
find $\dfrac{\partial u}{\partial x}$.

11. If
$$\begin{cases} x = u + vu^v, \\ y = v - uv^u, \end{cases}$$
find $\dfrac{\partial u}{\partial x}$.

12. If $\qquad u = x^2 + y^2 + z^2 \qquad$ and $\qquad z = xyt,$

explain all the meanings of $\dfrac{\partial u}{\partial x}$.

13. If
$$\begin{cases} z = f(x, y), \\ \phi(x, y) = 0, \end{cases}$$

show that
$$\frac{dz}{dx} = \frac{\dfrac{\partial z}{\partial x}\dfrac{\partial \phi}{\partial y} - \dfrac{\partial z}{\partial y}\dfrac{\partial \phi}{\partial x}}{\dfrac{\partial \phi}{\partial y}}.$$

14. If
$$u = f(x + \alpha t, y + \beta t),$$

show that
$$\frac{\partial u}{\partial t} = \alpha \frac{\partial u}{\partial x} + \beta \frac{\partial u}{\partial y},$$

$$\frac{\partial^2 u}{\partial t^2} = \alpha^2 \frac{\partial^2 u}{\partial x^2} + 2\alpha\beta \frac{\partial^2 u}{\partial x \partial y} + \beta^2 \frac{\partial^2 u}{\partial y^2},$$

and obtain the general formula for $\dfrac{\partial^n u}{\partial t^n}$.

15. If
$$u = f(y + ax) + \phi(y - ax),$$

show that
$$\frac{\partial^2 u}{\partial x^2} = a^2 \frac{\partial^2 u}{\partial y^2}.$$

16. If
$$u = f\left(\frac{y}{x}\right),$$

show that
$$x \frac{\partial u}{\partial x} + y \frac{\partial u}{\partial y} = 0.$$

17. Use the method of differentials to find $\dfrac{\partial u}{\partial x}$, $\dfrac{\partial u}{\partial y}$, and $\dfrac{\partial u}{\partial t}$, in terms of $f_1(\xi, \eta)$, $f_2(\xi, \eta)$, if
$$u = f(x + ut, y - ut).$$

18. If u is a function merely of the differences of the arguments $x_1, x_2, \cdots, x_n$, show that
$$\frac{\partial u}{\partial x_1} + \frac{\partial u}{\partial x_2} + \cdots + \frac{\partial u}{\partial x_n} = 0.$$

19. If $u = f(x, y)$ is homogeneous of order n, show that
$$x^2 \frac{\partial^2 u}{\partial x^2} + 2xy \frac{\partial^2 u}{\partial x \partial y} + y^2 \frac{\partial^2 u}{\partial y^2} = n(n-1)u.$$

20. Extend the theorem of Question 19 to the case of homogeneous functions of three variables.

21. Extend the theorem of Question 19 to the case of derivatives of the third order.

22. If u and v are two functions of x and y satisfying the relations:

$$\frac{\partial u}{\partial x}=\frac{\partial v}{\partial y}, \qquad \frac{\partial u}{\partial y}=-\frac{\partial v}{\partial x},$$

show that, on introducing polar coordinates:

$$x=r\cos\phi, \qquad y=r\sin\phi,$$

we have

$$\frac{\partial u}{\partial r}=\frac{1}{r}\frac{\partial v}{\partial \phi}, \qquad \frac{1}{r}\frac{\partial u}{\partial \phi}=-\frac{\partial v}{\partial r}.$$

23. Under the hypotheses of the preceding question, show that

$$\frac{\partial^2 u}{\partial r^2}+\frac{1}{r}\frac{\partial u}{\partial r}+\frac{1}{r^2}\frac{\partial^2 u}{\partial \phi^2}=0.$$

24. If

$$f(x, y)=0 \qquad \text{and} \qquad \phi(x, z)=0,$$

show that

$$\frac{\partial \phi}{\partial x}\frac{\partial f}{\partial y}\frac{dy}{dz}=\frac{\partial f}{\partial x}\frac{\partial \phi}{\partial z}.$$

25. If

$$\phi(p, v, t)=0,$$

show that

$$\frac{\partial p}{\partial t}\frac{\partial t}{\partial v}\frac{\partial v}{\partial p}=-1.$$

Explain the meaning of each of the partial derivatives.

26. If u is a function of x, y, z and x, y, z are connected by a single relation, is it true that

$$\frac{\partial u_{xy}}{\partial y}=\frac{\partial u_{xz}}{\partial z}\frac{\partial z_{xy}}{\partial y} \ ?$$

27. If $u=f(x, y)$ and $v=\phi(x, y)$ are two functions which satisfy the relations

$$\frac{\partial u}{\partial x}=\frac{\partial v}{\partial y}, \qquad \frac{\partial u}{\partial y}=-\frac{\partial v}{\partial x},$$

and if V is any third function, show that

$$\frac{\partial^2 V}{\partial x^2}+\frac{\partial^2 V}{\partial y^2}=\left[\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial u}{\partial y}\right)^2\right]\left(\frac{\partial^2 V}{\partial u^2}+\frac{\partial^2 V}{\partial v^2}\right).$$

28. If

$$x=r\sin\phi\cos\theta, \qquad y=r\sin\phi\sin\theta, \qquad z=r\cos\phi,$$

show that

$$\left(\frac{\partial V}{\partial x}\right)^2+\left(\frac{\partial V}{\partial y}\right)^2+\left(\frac{\partial V}{\partial z}\right)^2=\left(\frac{\partial V}{\partial r}\right)^2+\frac{1}{r^2}\left(\frac{\partial V}{\partial \phi}\right)^2+\frac{1}{r^2\sin^2\phi}\left(\frac{\partial V}{\partial \theta}\right)^2.$$

29. If

$$x=r\sin\phi\cos\theta, \qquad y=r\sin\phi\sin\theta, \qquad z=r\cos\phi,$$

show that

$$\frac{\partial^2 V}{\partial x^2}+\frac{\partial^2 V}{\partial y^2}+\frac{\partial^2 V}{\partial z^2}=\frac{\partial^2 V}{\partial r^2}+\frac{1}{r^2}\frac{\partial^2 V}{\partial \phi^2}+\frac{1}{r^2\sin^2\phi}\frac{\partial^2 V}{\partial \theta^2}+\frac{2}{r}\frac{\partial V}{\partial r}+\frac{\cot\phi}{r^2}\frac{\partial V}{\partial \phi}.$$

30. If

$$dU=\theta\, dS-p dv$$

is an exact differential (p. 356), and if S and v can be expressed as functions of the independent variables θ, p, show that

$$\frac{\partial \theta}{\partial v}=-\frac{\partial p}{\partial S}, \qquad \frac{\partial S}{\partial p}=-\frac{\partial v}{\partial \theta}.$$

State what the independent variables are in each differentiation.

31. Let $\quad x=f(t, u), \qquad y=\phi(t, u), \qquad z=\psi(t, u),$

and

$$\frac{\partial(x, y)}{\partial(t, u)}\neq 0.$$

If

$$z=\Psi(x, y)$$

represents the relation connecting x, y, z, show that

$$\frac{\partial z}{\partial x}=\frac{\partial(z, y)}{\partial(t, u)}\Big/\frac{\partial(x, y)}{\partial(t, u)}, \qquad \frac{\partial z}{\partial y}=\frac{\partial(x, z)}{\partial(t, u)}\Big/\frac{\partial(x, y)}{\partial(t, u)}.$$

32. If

$$u=f(x, y, z), \qquad v=\phi(x, y, z), \qquad w=\psi(x, y, z)$$

and $\qquad x=g(\lambda, \mu), \qquad y=h(\lambda, \mu), \qquad z=k(\lambda, \mu),$

show that

$$\frac{\partial(v, w)}{\partial(\lambda, \mu)}=\frac{\partial(v, w)}{\partial(y, z)}\frac{\partial(y, z)}{\partial(\lambda, \mu)}+\frac{\partial(v, w)}{\partial(z, x)}\frac{\partial(z, x)}{\partial(\lambda, \mu)}+\frac{\partial(v, w)}{\partial(x, y)}\frac{\partial(x, y)}{\partial(\lambda, \mu)}.$$

State the conditions of continuity which you assume.

33. Let

$$u=f(x, y), \qquad v=\phi(x, y),$$

$$\frac{\partial(u, v)}{\partial(x, y)}\neq 0,$$

the usual conditions of continuity being assumed in the neighborhood of (x_0, y_0). Show that the Jacobian is positive at (x_0, y_0) if a small circle about this point, when described in the counter-clockwise sense, goes over into a small oval about (u_0, v_0), likewise described in the counter-clockwise sense. But when the sense is reversed, the Jacobian is negative.

34. Let

$$u = f(x, y, z), \qquad v = \phi(x, y, z), \qquad w = \psi(x, y, z),$$

$$\frac{\partial(u, v, w)}{\partial(x, y, z)} \neq 0,$$

the usual conditions of continuity being assumed in the neighborhood of (x_0, y_0, z_0). Show that the Jacobian is positive at (x_0, y_0, z_0) if the positive directions of the curvilinear coordinates (u, v, w) are oriented there as the positive directions of the (x, y, z)-coordinates; otherwise, the Jacobian is negative.

35. If $F(u, v, x, y)$ and $\Phi(u, v, x, y)$ are two functions which satisfy the conditions of § 10, show that

$$\frac{\partial(u, v)}{\partial(x, y)} = -\frac{\dfrac{\partial(F, \Phi)}{\partial(x, y)}}{\dfrac{\partial(F, \Phi)}{\partial(u, v)}}.$$

Is the corresponding theorem true in the general case, $n = n$?

CHAPTER VI

APPLICATIONS TO THE GEOMETRY OF SPACE

1. Tangent Plane and Normal Line to a Surface. (*a*) *Explicit Form of the Equation of the Surface.* Let the equation of the surface be given in the explicit form,

(1) $$z = f(x, y).$$

Then the equation of the tangent plane at the point (x_0, y_0, z_0) is (cf. *Introduction to the Calculus*, Chap. XV, § 3):

(2) $$z - z_0 = \left(\frac{\partial z}{\partial x}\right)_0 (x - x_0) + \left(\frac{\partial z}{\partial y}\right)_0 (y - y_0).$$

The equation of the normal line at the same point is

(3) $$\frac{x - x_0}{\left(\frac{\partial z}{\partial x}\right)_0} = \frac{y - y_0}{\left(\frac{\partial z}{\partial y}\right)_0} = \frac{z - z_0}{-1}.$$

Finally, the direction cosines of the normal at an arbitrary point (x, y, z) of (1) satisfy the relations:

(4) $$\cos\alpha : \cos\beta : \cos\gamma = \frac{\partial z}{\partial x} : \frac{\partial z}{\partial y} : -1.$$

(*b*) *Implicit Form.* If the equation of the surface is given in the implicit form,

(5) $$F(x, y, z) = 0,$$

it follows then from (2) and the expressions for the partial derivatives, Chap. V, § 9, (8), that the equation of the tangent plane is

(6) $$\left(\frac{\partial F}{\partial x}\right)_0 (x - x_0) + \left(\frac{\partial F}{\partial y}\right)_0 (y - y_0) + \left(\frac{\partial F}{\partial z}\right)_0 (z - z_0) = 0.$$

For the normal line,

(7) $$\frac{x - x_0}{\left(\frac{\partial F}{\partial x}\right)_0} = \frac{y - y_0}{\left(\frac{\partial F}{\partial y}\right)_0} = \frac{z - z_0}{\left(\frac{\partial F}{\partial z}\right)_0},$$

and for the direction cosines of the normal at (x, y, z),

(8) $$\cos\alpha : \cos\beta : \cos\gamma = \frac{\partial F}{\partial x} : \frac{\partial F}{\partial y} : \frac{\partial F}{\partial z}.$$

(c) ***Parametric Form.*** Let the equation of the surface be given in the parametric form

(9) $$x = f(u, v), \qquad y = \phi(u, v), \qquad z = \psi(u, v),$$

where f, ϕ, ψ are continuous, together with their first derivatives, and where, moreover, at least one of the two-rowed determinants out of the matrix

(10) $$\left\| \begin{matrix} \frac{\partial f}{\partial u} & \frac{\partial \phi}{\partial u} & \frac{\partial \psi}{\partial u} \\ \frac{\partial f}{\partial v} & \frac{\partial \phi}{\partial v} & \frac{\partial \psi}{\partial v} \end{matrix} \right\|,$$

i.e. at least one of the Jacobians

(11) $$\frac{\partial(y, z)}{\partial(u, v)}, \qquad \frac{\partial(z, x)}{\partial(u, v)}, \qquad \frac{\partial(x, y)}{\partial(u, v)},$$

is different from 0 at the point (u_0, v_0) corresponding to (x_0, y_0, z_0). Then the equation of the tangent plane, as will presently be shown, is

(12) $$\left(\frac{\partial(y, z)}{\partial(u, v)}\right)_0 (x - x_0) + \left(\frac{\partial(z, x)}{\partial(u, v)}\right)_0 (y - y_0) + \left(\frac{\partial(x, y)}{\partial(u, v)}\right)_0 (z - z_0) = 0.$$

The equations of the normal line are

(13) $$\frac{x - x_0}{\left(\frac{\partial(y, z)}{\partial(u, v)}\right)_0} = \frac{y - y_0}{\left(\frac{\partial(z, x)}{\partial(u, v)}\right)_0} = \frac{z - z_0}{\left(\frac{\partial(x, y)}{\partial(u, v)}\right)_0}.$$

The direction cosines of the normal satisfy the relations:

(14) $$\cos\alpha : \cos\beta : \cos\gamma = \frac{\partial(y, z)}{\partial(u, v)} : \frac{\partial(z, x)}{\partial(u, v)} : \frac{\partial(x, y)}{\partial(u, v)}.$$

The proof is given at once by Chap. V, p. 150, Ex. 31, from which it follows that (14) is true.

EXERCISES

1. On writing the equations of the sphere $x^2 + y^2 + z^2 = a^2$ in the parametric form:

$$x = a \sin\phi \cos\theta, \qquad y = a \sin\phi \sin\theta, \qquad z = a \cos\phi,$$

show that the normal at an arbitrary point goes through the centre. Are there exceptions?

2. Express the equations of the torus parametrically, and show that the normal at any point goes through the axis. Are there exceptions?

2. Analytic Representation of Space Curves. A curve in space may be given analytically by expressing its coordinates

(*a*) as functions of a parameter:

(1) $$x = f(t), \qquad y = \phi(t), \qquad z = \psi(t);$$

(*b*) as the intersection of two cylinders:

(2) $$y = \phi(x), \qquad z = \psi(x);$$

(*c*) as the intersection of two arbitrary surfaces:

(3) $$F(x, y, z) = 0, \qquad \Phi(x, y, z) = 0.$$

A familiar example of (*a*) in the case of plane curves is the cycloid; also the circle. In the case of space curves we have the helix:

(4) $$x = a\cos\theta, \qquad y = a\sin\theta, \qquad z = b\theta.$$

This curve winds round the cylinder $x^2 + y^2 = a^2$, its steepness always keeping the same. It is the curve of the thread of a machine screw, *i.e.* of a screw that does not taper. Again, if a body is moving under a given law of force, the coordinates of its centre of gravity are functions of the time, and we may think of these as expressed in the form (*a*). But the student must not regard it as essential that we find a simple geometrical or mechanical interpretation for t in (*a*). Thus if we write arbitrarily:

(5) $$x = \log t, \qquad y = \sin t, \qquad z = \frac{t}{\sqrt[3]{1 + t^2}},$$

we get a perfectly good curve, t entering purely analytically.

In particular, we can always choose as the parameter t in (*a*) the length of the arc of the curve, measured from an arbitrary point:

(6) $$x = f(s), \qquad y = \phi(s), \qquad z = \psi(s).$$

The form (*b*) may be regarded as a special case under (*a*), namely, that in which
$$x = t.$$

On the other hand, it is a special case under (*c*).

Restrictions on the Functions. It is natural enough to require, in Case (*b*), that the functions $\phi(x)$ and $\psi(x)$ be continuous, together with their first derivatives; and in Case (*c*), that the functions $F(x, y, z)$ and $\Phi(x, y, z)$ fulfil the conditions of the existence theorem of Chap. V, § 12. In Case (*a*), however, the continuity of the functions $f(t)$, etc., together with that of their first derivatives, is not enough to insure a curve; for, this condition is satisfied when all three functions are constants, and then equations (1) represent a point.

Again, the plane curve

$$(7)\qquad x = t^2, \qquad y = t^3, \qquad i.e. \qquad y^2 = x^3,$$

has a cusp at the origin, and yet the derivatives of the functions of t are continuous there.

It is sufficient, in order to avoid all these difficulties, to demand that the first derivatives of the functions of t, — namely, $f'(t)$, $\phi'(t)$, $\psi'(t)$, — *never vanish simultaneously*, or that

$$(8)\qquad 0 < f'(t)^2 + \phi'(t)^2 + \psi'(t)^2,$$

and this condition shall henceforth be imposed in general. If, at a particular point of a curve, the condition is violated, such a point will usually be a singular point, as in the case of the curve (7) at the point $t = 0$. But this is not always true, as is shown by the example

$$(9)\qquad x = t^2, \qquad y = t^2, \qquad i.e. \qquad x = y.$$

Here, both derivatives vanish at the origin, but this point is in no wise peculiar.

To sum up, then, the condition (8) is *sufficient*, but *not necessary*. The proof that it is sufficient lies in the fact that, since at least one derivative, as $f'(t)$, is not 0, it is possible to solve for t. Here

$$x_0 = f(t_0), \qquad f'(t_0) \neq 0.$$

Hence, by Chap. V, § 12, p. 135,

$$t = \omega(x),$$

where $\omega(x_0) = t_0$ and $\omega(x)$ is continuous, together with its first derivative, at the point $x = x_0$. On substituting this value for t in the last two of the equations (1), these go over into the form (2), and hence we have a curve, q. e. d.

3. The Direction Cosines and the Arc. To find the direction cosines of the tangent to a space curve at a point P: (x_0, y_0, z_0), pass a secant through P and a neighboring point P': $(x_0 + \Delta x, y_0 + \Delta y, z_0 + \Delta z)$. The direction cosines of the secant are:

$$\cos \alpha' = \frac{\Delta x}{\overline{PP'}}, \qquad \cos \beta' = \frac{\Delta y}{\overline{PP'}}, \qquad \cos \gamma' = \frac{\Delta z}{\overline{PP'}},$$

and hence, for the tangent,

$$\cos \alpha = \lim_{\overline{PP'} \doteq 0} \frac{\Delta x}{\overline{PP'}} = \lim_{\overline{PP'} \doteq 0} \left(\frac{\Delta x}{\Delta s} \cdot \frac{\Delta s}{\overline{PP'}}\right) = D_s x,$$

with similar formulas for $\cos\beta$, $\cos\gamma$. Thus

(10) $$\cos\alpha = \frac{dx}{ds}, \qquad \cos\beta = \frac{dy}{ds}, \qquad \cos\gamma = \frac{dz}{ds}.$$

Here the tangent is thought of as drawn in the direction in which s is increasing. If it is drawn in the opposite direction, the minus sign must precede each derivative.

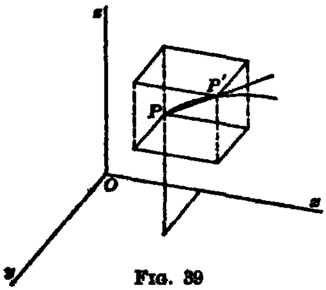


FIG. 39

From (10) it follows at once that

(11) $$ds^2 = dx^2 + dy^2 + dz^2.$$

This important formula can be proven directly from the relation

$$\overline{PP'}^2 = \Delta x^2 + \Delta y^2 + \Delta z^2.$$

If we assume the form (a),

$$ds^2 = [f'(t)^2 + \phi'(t)^2 + \psi'(t)^2]\,dt^2$$

and

(12) $$\begin{cases} \cos\alpha = \dfrac{f'(t)}{\sqrt{f'(t)^2 + \phi'(t)^2 + \psi'(t)^2}}, \\ \cos\beta = \dfrac{\phi'(t)}{\sqrt{f'(t)^2 + \phi'(t)^2 + \psi'(t)^2}}, \\ \cos\gamma = \dfrac{\psi'(t)}{\sqrt{f'(t)^2 + \phi'(t)^2 + \psi'(t)^2}}. \end{cases}$$

(13) $$s = \int_{t_0}^{t_1} \sqrt{f'(t)^2 + \phi'(t)^2 + \psi'(t)^2}\,dt.$$

Applying these results to (2), we get

(14) $$\cos\alpha = \frac{1}{\sqrt{1 + \frac{dy^2}{dx^2} + \frac{dz^2}{dx^2}}}, \quad \cos\beta = \frac{\frac{dy}{dx}}{\sqrt{1 + \frac{dy^2}{dx^2} + \frac{dz^2}{dx^2}}}, \quad \cos\gamma = \frac{\frac{dz}{dx}}{\sqrt{1 + \frac{dy^2}{dx^2} + \frac{dz^2}{dx^2}}}.$$

(15) $$s = \int_{x_0}^{x_1} \sqrt{1 + \frac{dy^2}{dx^2} + \frac{dz^2}{dx^2}}\,dx.$$

The direction cosines in Case (c) are obtained in § 4.

4. Equations of Tangent Line and Normal Plane. For the tangent line we have, in Case (a):

$$(16)\qquad \frac{x-x_0}{f'(t_0)}=\frac{y-y_0}{\phi'(t_0)}=\frac{z-z_0}{\psi'(t_0)};$$

and in (b):

$$(17)\qquad y-y_0=\left(\frac{dy}{dx}\right)_0(x-x_0),\qquad z-z_0=\left(\frac{dz}{dx}\right)_0(x-x_0).$$

The normal plane is given in (a) by

$$(18)\qquad f'(t_0)(x-x_0)+\phi'(t_0)(y-y_0)+\psi'(t_0)(z-z_0)=0;$$

and in (b) by

$$(19)\qquad x-x_0+\left(\frac{dy}{dx}\right)_0(y-y_0)+\left(\frac{dz}{dx}\right)_0(z-z_0)=0.$$

On the other hand, the tangent line in Case (c) may be obtained most simply as the intersection of the tangent planes to the surfaces at the point in question:

$$(20)\qquad \begin{cases}\left(\dfrac{\partial F}{\partial x}\right)_0(x-x_0)+\left(\dfrac{\partial F}{\partial y}\right)_0(y-y_0)+\left(\dfrac{\partial F}{\partial z}\right)_0(z-z_0)=0,\\[2ex] \left(\dfrac{\partial \Phi}{\partial x}\right)_0(x-x_0)+\left(\dfrac{\partial \Phi}{\partial y}\right)_0(y-y_0)+\left(\dfrac{\partial \Phi}{\partial z}\right)_0(z-z_0)=0.\end{cases}$$

These equations may be thrown into the equivalent form:

$$(21)\qquad \frac{x-x_0}{\begin{vmatrix}F_y & F_z\\ \Phi_y & \Phi_z\end{vmatrix}_0}=\frac{y-y_0}{\begin{vmatrix}F_z & F_x\\ \Phi_z & \Phi_x\end{vmatrix}_0}=\frac{z-z_0}{\begin{vmatrix}F_x & F_y\\ \Phi_x & \Phi_y\end{vmatrix}_0}.$$

Hence we see that the direction cosines of the tangent line to the curve of intersection of the surfaces (3) are given at (x, y, z) by the proportion:

$$(22)\qquad \cos\alpha:\cos\beta:\cos\gamma=\begin{vmatrix}F_y & F_z\\ \Phi_y & \Phi_z\end{vmatrix}:\begin{vmatrix}F_z & F_x\\ \Phi_z & \Phi_x\end{vmatrix}:\begin{vmatrix}F_x & F_y\\ \Phi_x & \Phi_y\end{vmatrix}.$$

The equation of the normal plane and the integral which represents the arc, s, can now be written down.

EXERCISES

Find the equations of the tangent line and the normal plane to each of the following space curves:

1. The helix (4).

2. The curve (5).

3. The curve: $y^2 = 2mx, \quad z^2 = m - x.$

4. The curve: $2x^2 + 3y^2 + z^2 = 9, \quad z^2 = 3x^2 + y^2,$

at the point $(1, -1, 2)$.

5. Find the angle that the tangent line in the preceding question makes with the axis of x.

6. Compute the length of the arc of the helix:

$$x = \cos\theta, \qquad y = \sin\theta, \qquad 5z = \theta,$$

when it has made one complete turn around the cylinder.

7. How steep is the helix in the preceding question?

8. Show that the condition that the surfaces (3) cut orthogonally is that

$$\frac{\partial F}{\partial x}\frac{\partial \Phi}{\partial x} + \frac{\partial F}{\partial y}\frac{\partial \Phi}{\partial y} + \frac{\partial F}{\partial z}\frac{\partial \Phi}{\partial z} = 0. \tag{23}$$

9. What is the condition that the three surfaces:

$$F(x, y, z) = 0, \qquad \Phi(x, y, z) = 0, \qquad \Psi(x, y, z) = 0,$$

intersecting at the point (x_0, y_0, z_0), be tangent to one and the same line there?

10. The surfaces

$$x^2 + y^2 + z^2 = 3, \qquad xyz = 1, \qquad z = xy,$$

all go through the point (1, 1, 1). Find the angles at which they intersect there.

11. Obtain the condition that the surface (1), § 1, and the curve (1), § 2 meet at right angles.

12. Find the direction angles of the curve

$$x = t^2, \qquad y = t^3, \qquad z = t^4$$

in the point (1, 1, 1).

13. Find the direction angles of the curve

$$xyz = 1, \qquad y^2 = x$$

in the point (1, 1, 1).

14. Find all the points in which the curve

$$x = t^2, \qquad y = t^3, \qquad z = t^4$$

meets the surface

$$z^2 = x + 2y - 2,$$

and show that, when it meets the surface, it is tangent to it.

15. Show that the surfaces

$$xyz = 1, \qquad \frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1,$$

in general never cut orthogonally; but that, if

$$\frac{1}{a^2} + \frac{1}{b^2} - \frac{1}{c^2} = 0,$$

they cut orthogonally along their whole line of intersection.

16. When will the spheres

$$x^2 + y^2 + z^2 = 1, \qquad (x-a)^2 + (y-b)^2 + (z-c)^2 = 1$$

cut orthogonally?

17. Two space curves have their equations written in the form (6). They intersect at a point P. Show that the angle ϵ between them at P is given by the equation:

$$\cos\epsilon = x_1'x_2' + y_1'y_2' + z_1'z_2', \qquad x_1' = \frac{dx_1}{ds}, \text{ etc.}$$

18. The ellipsoid: $x^2 + 3y^2 + 2z^2 = 9$ and the sphere: $x^2 + y^2 + z^2 = 6$ intersect in the point (2, 1, 1). Find the angle between their tangent planes at this point.

19. Let a surface be given in the parametric form (9), § 1, and let a curve in the (u, v)-plane also be given in parametric form:

$$u = \chi(t), \qquad v = \omega(t).$$

Show that, if $\chi(t)$ and $\omega(t)$ be substituted respectively for u and v in the functions $f(u, v)$, $\phi(u, v)$, $\psi(u, v)$, the new equations (9) represent a space curve.

20. Prove that all the space curves of the preceding question which pass through a given point of the surface are perpendicular to one and the same right line, and that the direction components of this line are

$$\frac{\partial(y, z)}{\partial(u, v)}, \qquad \frac{\partial(z, x)}{\partial(u, v)}, \qquad \frac{\partial(x, y)}{\partial(u, v)}$$

respectively, formed for this point.

21. If a surface is given in the parametric form (9), § 1, show that the differential of arc of a curve on the surface is given by the equation:

$$ds^2 = E\,du^2 + 2\,F\,du\,dv + G\,dv^2,$$

where

$$E = \left(\frac{\partial x}{\partial u}\right)^2 + \left(\frac{\partial y}{\partial u}\right)^2 + \left(\frac{\partial z}{\partial u}\right)^2,$$

$$F = \frac{\partial x}{\partial u}\frac{\partial x}{\partial v} + \frac{\partial y}{\partial u}\frac{\partial y}{\partial v} + \frac{\partial z}{\partial u}\frac{\partial z}{\partial v},$$

$$G = \left(\frac{\partial x}{\partial v}\right)^2 + \left(\frac{\partial y}{\partial v}\right)^2 + \left(\frac{\partial z}{\partial v}\right)^2.$$

In the above equation for ds, what is the independent variable?

5. Osculating Plane. Let $P: (x_0, y_0, z_0)$ be an arbitrary point of a space curve, (1), § 2, and pass a plane

(24) $$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0$$

through P. Then the distance D of a neighboring point

$$P': \qquad x = f(t_0 + h), \qquad y = \phi(t_0 + h), \qquad z = \psi(t_0 + h)$$

of the curve from this plane will be in general an infinitesimal of the first order with reference to $\overset{\frown}{PP'}$ as principal infinitesimal. For

$$\pm D = \frac{A(x - x_0) + B(y - y_0) + C(z - z_0)}{\sqrt{A^2 + B^2 + C^2}},$$

where x, y, z are the coordinates of P'. Hence

$$\pm D = \frac{A[f(t_0 + h) - f(t_0)] + B[\phi(t_0 + h) - \phi(t_0)] + \text{etc.}}{\sqrt{A^2 + B^2 + C^2}}.$$

Apply Taylor's Theorem with the Remainder (*Introduction to the Calculus*, p. 430) to each bracket:

$$f(t_0 + h) - f(t_0) = hf'(t_0) + \frac{h^2}{2}f''(t_0 + \theta h), \qquad 0 < \theta < 1,$$

etc.,

and set $\sqrt{A^2 + B^2 + C^2} = \Delta$. We thus obtain the following expression for D:

$$\pm D = h[Af'(t_0) + B\phi'(t_0) + C\psi'(t_0)]/\Delta$$

$$+ \frac{h^2}{2}[Af''(t_0 + \theta h) + B\phi''(t_0 + \theta_1 h) + C\psi''(t_0 + \theta_2 h)]/\Delta.$$

Hence $$\lim_{P' \doteq P} \frac{\pm D}{h} = \frac{Af'(t_0) + B\phi'(t_0) + C\psi'(t_0)}{\Delta},$$

and this will not $= 0$ if A, B, C are chosen at random, since $f'(t_0)$, $\phi'(t_0)$, $\psi'(t_0)$ cannot all vanish simultaneously, unless perchance at an exceptional point. On the other hand, $\widetilde{PP'} = \Delta s$ and $h = \Delta t$ are infinitesimals of the same order, since

$$\lim_{\Delta t \doteq 0} \frac{\Delta s}{\Delta t} = D_t s = \sqrt{f'(t_0)^2 + \phi'(t_0)^2 + \psi'(t_0)^2} \neq 0.$$

Thus the above statement is proven.

If, however, A, B, and C are so chosen that

$$(25) \qquad Af'(t_0) + B\phi'(t_0) + C\psi'(t_0) = 0,$$

then $\lim \pm D/h = 0$ and

$$\lim_{P' \doteq P} \frac{\pm D}{h^2} = \frac{Af''(t_0) + B\phi''(t_0) + C\psi''(t_0)}{2\Delta}.$$

Now (25) is precisely the condition that the tangent line to (1), § 2, be perpendicular to the normal to the plane (24), and hence the tangent will lie in this plane; *i.e.* the plane (24) is here tangent to the curve, and D becomes now in general an infinitesimal of the second order. But if A, B, and C are furthermore subject to the restriction that

$$(26) \qquad Af''(t_0) + B\phi''(t_0) + C\psi''(t_0) = 0,$$

then even $\lim \pm D/h^2 = 0$ and D becomes an infinitesimal of still higher order; — of the *third* order, as is readily shown, if

$$Af'''(t_0) + B\phi'''(t_0) + C\psi'''(t_0) \neq 0.$$

Equations (25) and (26) serve in general to define the ratios of the coefficients A, B, C uniquely. The latter may, therefore, be eliminated from (24), (25), and (26), and thus we obtain the equation of the *osculating plane:*

$$(27) \qquad \begin{vmatrix} x - x_0 & y - y_0 & z - z_0 \\ f'(t_0) & \phi'(t_0) & \psi'(t_0) \\ f''(t_0) & \phi''(t_0) & \psi''(t_0) \end{vmatrix} = 0.$$

The osculating plane as thus defined is a tangent plane having contact of higher order than one of the tangent planes taken at random. There is in general only one osculating plane at a given point. But in the case of a straight line all tangent planes osculate. Again, if $f''(t_0) = \phi''(t_0) = \psi''(t_0) = 0$, the same is true. And, generally, all tangent planes osculate whenever all the two-rowed determinants formed from the elements of the last two rows of (27) vanish. The osculating plane cuts the curve in general at the point of tangency; for the numerator of the expression for $\pm D$ changes sign when h passes through the value 0.

It is easy to make a simple model that will show the osculating plane approximately. Wind a piece of soft iron wire round a broom handle, thus making a helix, and then cut out an inch of the wire and lay it down on a table. The piece will look almost like a plane curve in the plane of the table, and the latter will be approximately the osculating plane.

A second experiment that can be made with the first model described — I mean, the longer wire — is, to hold it up and sight along the tangent at an arbitrary point P, thus projecting the wire on the wall. The projection will be seen to have a cusp at the point which corresponds to P, — and this, no matter what point P be chosen.

The normal line to a space curve, drawn in the osculating plane, is called the *principal normal*. The *centre of curvature* lies on this line, the *radius of curvature* being obtained by projecting the curve orthogonally on the osculating plane and taking the radius of curvature of this projection. The line through P perpendicular to both the tangent and the principal normal is called the *bi-normal*.

If a body move under the action of any forces, the vector acceleration of its centre of gravity always lies in the osculating plane of the path.

When the equation of the curve is given in the form (2), § 2, the equation (27) becomes:

$$(28)\quad \left(\frac{dz}{dx}\frac{d^2y}{dx^2}-\frac{dy}{dx}\frac{d^2z}{dx^2}\right)_0(x-x_0)+\left(\frac{d^2z}{dx^2}\right)_0(y-y_0)-\left(\frac{d^2y}{dx^2}\right)_0(z-z_0)=0.$$

EXERCISES

1. Find the equation of the osculating plane of the curve (5), § 2, at the point $t=\pi$.

2. Find the equation of the osculating plane of the curve of intersection of the cylinders:

$$x^2+y^2=a^2, \qquad x^2+z^2=a^2,$$

and interpret the result.

Suggestion. Express x, y, z in terms of t, as for example:

$$x=a\cos t, \qquad y=a\sin t, \qquad z=a\sin t.$$

3. Show that the centre of curvature of a helix lies on the radius of the cylinder produced.

4. Show that the osculating plane of the curve

$$y = x^2, \qquad z^2 = 1 - y$$

at the point (0, 0, 1) has contact of higher order than the second.

5. Prove that, in the case of a plane curve, the osculating plane is the plane of the curve.

6. Show that, if the tangent to a space curve at a given point P be taken as the axis of x, the principal normal as the axis of y, and the bi-normal as the axis of z, equation (27) reduces to $z = 0$.

7. Work out the equation of the osculating plane when the curve is given by equations (3), § 2.

8. Prove that, if the osculating plane to a space curve is parallel at every point to a fixed plane, the curve is a plane curve.

9. Show that, if all the tangent planes to a curve at an arbitrary point osculate, *i.e.* if, no matter where the point P be taken on the curve, every tangent plane at P has contact of higher order than is in general the case, then the curve is a straight line.

6. Confocal Quadrics and Orthogonal Systems.* Consider the family of surfaces:

(1) $$\frac{x^2}{a^2+\lambda} + \frac{y^2}{b^2+\lambda} + \frac{z^2}{c^2+\lambda} = 1, \qquad a > b > c > 0,$$

where λ is a parameter taking on different values. Each surface of the family is symmetric with regard to each of the coordinate planes. We may, therefore, confine ourselves to the first octant.

If $\lambda > -c^2$, we have an ellipsoid, which for large positive values of λ resembles a huge sphere. As λ decreases, the surface contracts, and as λ approaches $-c^2$, the ellipsoid, whose equation can be thrown into the form:

$$z^2 = (c^2+\lambda)\left(1 - \frac{x^2}{a^2+\lambda} - \frac{y^2}{b^2+\lambda}\right),$$

flattens down toward the plane $z = 0$ as its limit, — more precisely, toward the surface of the ellipse

$$\frac{x^2}{a^2-c^2} + \frac{y^2}{b^2-c^2} = 1, \qquad z = 0.$$

* Cf. *Analytic Geometry*, p. 590.

In so doing, it sweeps out the whole first octant just once, as we shall presently show analytically.

Let λ continue to decrease. We then get the family:

$$(2) \qquad \frac{x^2}{a^2+\mu}+\frac{y^2}{b^2+\mu}-\frac{z^2}{-(c^2+\mu)}=1, \quad -b^2<\mu<-c^2.$$

These are hyperboloids of one nappe, opening out on the axis of z, and they rise from coincidence with the plane $z=0$ for values of μ just under $-c^2$, sweep out the whole octant, and flatten out again toward the plane $y=0$ as their limit when μ approaches $-b^2$.

Finally, let λ trace out the interval from $-b^2$ to $-a^2$. We then get the hyperboloids of two nappes, cutting the axis of x:

$$(3) \qquad \frac{x^2}{a^2+\nu}-\frac{y^2}{-(b^2+\nu)}-\frac{z^2}{-(c^2+\nu)}=1, \quad -a^2<\nu<-b^2.$$

These start from coincidence with the plane $y=0$ when ν is near $-b^2$, sweep out the octant, and approach the plane $x=0$ as ν approaches $-a^2$.

THEOREM 1. *Through each point of the first octant passes one surface of each family, and only one.*

Let $P:(x, y, z)$, be an arbitrary point of this octant. Then $x>0$, $y>0$, $z>0$. Hold x, y, z fast and consider the function of λ:

$$f(\lambda)=\frac{x^2}{a^2+\lambda}+\frac{y^2}{b^2+\lambda}+\frac{z^2}{c^2+\lambda}-1.$$

The function is continuous except when $\lambda=-c^2, \quad -b^2, \quad \text{or} \quad -a^2$. In the interval $-c^2<\lambda<+\infty$ we have*

$$f(+\infty)=-1, \qquad \lim_{\lambda\doteq -c^2+} f(\lambda)=+\infty.$$

Hence the curve

$$u=f(\lambda)$$

crosses the axis of abscissas *at least* once in this interval.

On the other hand

$$f'(\lambda)=-\frac{x^2}{(a^2+\lambda)^2}-\frac{y^2}{(b^2+\lambda)^2}-\frac{z^2}{(c^2+\lambda)^2}<0.$$

Hence $f(\lambda)$ always increases as λ decreases, and so the curve cuts the axis *only once* in this interval. We see, therefore, that one and only one ellipsoid passes through the point P.

* By the notation $\lim\limits_{x\doteq a+} F(x)$ is meant the limit when x approaches a from above. Similarly, $\lim\limits_{x\doteq a-} F(x)$ means the limit when x approaches a from below.

Similar reasoning applied to the intervals $(-b^2, -c^2)$ and $(-a^2, -b^2)$ shows that one and only one hyperbola of one nappe, and one and only one hyperbola of two nappes pass through P.

Theorem 2. *The three quadrics through P intersect at right angles there.*

The condition that two surfaces intersect at right angles is given by (23), §4, Ex. 8. Applying this theorem to the surfaces (1) and (2), we wish to show that

$$\frac{2x}{a^2+\lambda}\frac{2x}{a^2+\mu}+\frac{2y}{b^2+\lambda}\frac{2y}{b^2+\mu}+\frac{2z}{c^2+\lambda}\frac{2z}{c^2+\mu}=0.$$

Now subtract equation (2) from equation (1):

$$(\mu-\lambda)\left[\frac{x^2}{(a^2+\lambda)(a^2+\mu)}+\frac{y^2}{(b^2+\lambda)(b^2+\mu)}+\frac{z^2}{(c^2+\lambda)(c^2+\mu)}\right]=0.$$

Since $\mu-\lambda\neq 0$, this proves the theorem.

The three systems of surfaces that we have here investigated are analogous to the three families of planes in Cartesian coordinates, to the spheres, planes, and cones in spherical polar coordinates, and to the planes, cylinders, and planes in cylindrical polar coordinates. They form what is called an *orthogonal system* of surfaces, and enable us to assign to the points of the first octant the coordinates (λ, μ, ν), where

$$-c^2<\lambda<+\infty,\qquad -b^2<\mu<-c^2,\qquad -a^2<\nu<-b^2.$$

EXERCISES

1. Let $\quad \lambda=f(x, y, z),\ \ \mu=\phi(x, y, z),\ \ \nu=\psi(x, y, z)$

be the equations of three orthogonal families; the functions f, ϕ, ψ having continuous first partial derivatives, and their Jacobian not vanishing. Show that

$$\frac{\partial\lambda}{\partial x}\frac{\partial\mu}{\partial x}+\frac{\partial\lambda}{\partial y}\frac{\partial\mu}{\partial y}+\frac{\partial\lambda}{\partial z}\frac{\partial\mu}{\partial z}=0,$$

with two further equations obtained by advancing the letters λ, μ, ν cyclically.

2. Show that

$$\frac{\partial x}{\partial\lambda}\frac{\partial x}{\partial\mu}+\frac{\partial y}{\partial\lambda}\frac{\partial y}{\partial\mu}+\frac{\partial z}{\partial\lambda}\frac{\partial z}{\partial\mu}=0,$$

with two further similar relations.

3. Let a curve be drawn through a point of the region, and let its equations be

$$x = g(t), \qquad y = h(t), \qquad z = k(t),$$

where g, h, k have continuous first derivatives, not all zero. Then

$$\frac{ds^2}{dt^2} = H_1 \frac{d\lambda^2}{dt^2} + H_2 \frac{d\mu^2}{dt^2} + H_3 \frac{d\nu^2}{dt^2},$$

where

$$H_1 = \frac{\partial x^2}{\partial \lambda^2} + \frac{\partial y^2}{\partial \lambda^2} + \frac{\partial z^2}{\partial \lambda^2},$$

$$H_2 = \frac{\partial x^2}{\partial \mu^2} + \frac{\partial y^2}{\partial \mu^2} + \frac{\partial z^2}{\partial \mu^2},$$

$$H_3 = \frac{\partial x^2}{\partial \nu^2} + \frac{\partial y^2}{\partial \nu^2} + \frac{\partial z^2}{\partial \nu^2}.$$

It is in this sense that the equation

$$ds^2 = H_1\, d\lambda^2 + H_2\, d\mu^2 + H_3\, d\nu^2$$

is to be understood.

4. By making use of the theorem for the multiplication of deter minants in the form given in Chap. XII, § 2, Ex. 5, show that

$$J^2 = H_1 H_2 H_3, \qquad \text{where} \qquad J = \frac{\partial(x, y, z)}{\partial(\lambda, \mu, \nu)}.$$

5. Prove the relations:

$$H_1 \frac{\partial \lambda}{\partial x} = \frac{\partial x}{\partial \lambda}, \qquad H_2 \frac{\partial \mu}{\partial x} = \frac{\partial x}{\partial \mu}, \qquad H_3 \frac{\partial \nu}{\partial x} = \frac{\partial x}{\partial \nu};$$

$$H_1 \frac{\partial \lambda}{\partial y} = \frac{\partial y}{\partial \lambda}, \qquad H_2 \frac{\partial \mu}{\partial y} = \frac{\partial y}{\partial \mu}, \qquad H_3 \frac{\partial \nu}{\partial y} = \frac{\partial y}{\partial \nu};$$

$$H_1 \frac{\partial \lambda}{\partial z} = \frac{\partial z}{\partial \lambda}, \qquad H_2 \frac{\partial \mu}{\partial z} = \frac{\partial z}{\partial \mu}, \qquad H_3 \frac{\partial \nu}{\partial z} = \frac{\partial z}{\partial \nu}.$$

Suggestion: Start with the equations

$$\frac{\partial x}{\partial \lambda}\frac{\partial \lambda}{\partial x} + \frac{\partial x}{\partial \mu}\frac{\partial \mu}{\partial x} + \frac{\partial x}{\partial \nu}\frac{\partial \nu}{\partial x} = 1,$$

$$\frac{\partial y}{\partial \lambda}\frac{\partial \lambda}{\partial x} + \frac{\partial y}{\partial \mu}\frac{\partial \mu}{\partial x} + \frac{\partial y}{\partial \nu}\frac{\partial \nu}{\partial x} = 0,$$

$$\frac{\partial z}{\partial \lambda}\frac{\partial \lambda}{\partial x} + \frac{\partial z}{\partial \mu}\frac{\partial \mu}{\partial x} + \frac{\partial z}{\partial \nu}\frac{\partial \nu}{\partial x} = 0;$$

multiply them respectively by $\partial x/\partial \lambda$, $\partial y/\partial \lambda$, $\partial z/\partial \lambda$, and add.

6. Show that

$$\Delta_1\lambda = \frac{1}{H_1}, \qquad \Delta_1\mu = \frac{1}{H_2}, \qquad \Delta_1\nu = \frac{1}{H_3},$$

where Δ_1 denotes the first differential operator,

$$\Delta_1 = \left(\frac{\partial}{\partial x}\right)^2 + \left(\frac{\partial}{\partial y}\right)^2 + \left(\frac{\partial}{\partial z}\right)^2.$$

7. Prove that, if u be any function of x, y, z, continuous together with its first partial derivatives, then

$$\frac{\partial u^2}{\partial x^2} + \frac{\partial u^2}{\partial y^2} + \frac{\partial u^2}{\partial z^2} = \frac{1}{H_1}\frac{\partial u^2}{\partial \lambda^2} + \frac{1}{H_2}\frac{\partial u^2}{\partial \mu^2} + \frac{1}{H_3}\frac{\partial u^2}{\partial \nu^2},$$

or

$$\Delta_1 u = \Delta_1\lambda\left(\frac{\partial u}{\partial \lambda}\right)^2 + \Delta_1\mu\left(\frac{\partial u}{\partial \mu}\right)^2 + \Delta_1\nu\left(\frac{\partial u}{\partial \nu}\right)^2.$$

7. Curves on the Sphere, Cylinder, and Cone. In order to study the properties of curves drawn on the surface of a sphere, we introduce as coordinates of the points of the surface the longitude θ and the latitude ϕ. Any curve can then be represented by the equation

$$(1) \qquad F(\theta, \phi) = 0.$$

To determine the angle ω between this curve and a parallel of latitude, draw the meridians and the parallels of latitude through an arbitrary point $P:(\theta_0, \phi_0)$ and a neighboring point $P':(\theta_0 + \Delta\theta, \phi_0 + \Delta\phi)$ of this curve. We thus obtain a small curvilinear rectangle, of which the arc PP' is the diagonal. We wish to determine the angle

$$\omega = \angle MPP'.$$

Now consider, beside the curvilinear right triangle MPP', a rectilinear right triangle whose hypotenuse is the chord PP' and one of whose legs is the perpendicular PM_1 let fall from P on the meridian plane through P'. The angle

$$\omega' = \angle M_1PP'$$

of this triangle evidently approaches ω as its limit when P' approaches P. We have:

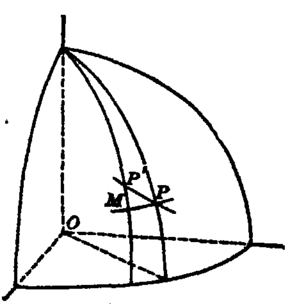


FIG. 40

$$\tan\omega' = \frac{M_1P'}{PM_1}.$$

It is clear that PM_1 differs from $\overset{\frown}{PM} = a \cos \phi_0 \Delta\theta$ by an infinitesimal of higher order, and likewise M_1P' differs from $\overset{\frown}{MP'} = a\,\Delta\phi$ by an infinitesimal of higher order. Hence, by the theorem of the *Introduction to the Calculus*, p. 90, which says that the limit of the ratio of two infinitesimals is unchanged if either or both the infinitesimals be replaced by others which are equivalent respectively to these, we obtain:

$$\lim_{P' \doteq P} \tan \omega' = \lim_{P' \doteq P} \frac{M_1P'}{PM_1} = \lim_{\Delta\theta \doteq 0} \frac{a\,\Delta\phi}{a \cos \phi_0 \Delta\theta},$$

$$\tan \omega = \frac{1}{\cos \phi_0} D_\theta \phi,$$

or, dropping the subscript:

$$\tan \omega = \frac{1}{\cos \phi} \frac{d\phi}{d\theta}. \tag{2}$$

In order to obtain the differential of the arc of the curve (1) we write down the Pythagorean Theorem for the triangle PM_1P':

$$\overline{PP'}^2 = \overline{PM_1}^2 + \overline{M_1P'}^2,$$

divide through by $\Delta\theta^2$, and then let $\Delta\theta$ approach 0 as its limit. Since the chord PP' differs from the arc Δs by an infinitesimal of higher order, we have:

$$\lim_{P' \doteq P} \left(\frac{PP'}{\Delta\theta}\right)^2 = \lim_{P' \doteq P} \left(\frac{\Delta s}{\Delta\theta}\right)^2 = a^2 \cos^2 \phi_0 + a^2 \lim_{P' \doteq P} \left(\frac{\Delta\phi}{\Delta\theta}\right)^2,$$

$$(D_\theta s)^2 = a^2 \cos^2 \phi + a^2 (D_\theta \phi)^2,$$

$$ds^2 = a^2[\cos^2 \phi \, d\theta^2 + d\phi^2]. \tag{3}$$

Rhumb lines. A *rhumb line* or *loxodrome* is the path of a ship that sails without altering her course, *i.e.* a curve that cuts the meridians always at one and the same angle. If we denote the complement of this angle by ω, then we have from (2) for the determination of the curve:

$$\frac{d\phi}{\cos \phi} = d\theta \tan \omega.$$

$$\theta \tan \omega = \int \frac{d\phi}{\cos \phi} = \log \tan\left(\frac{\phi}{2} + \frac{\pi}{4}\right) + C. \tag{4}$$

This is the equation of an equiangular spiral on the sphere, which winds round each of the poles an infinite number of times.

EXERCISES

1. Show that the total length of a rhumb line on the sphere is finite.

2. The Cartesian coordinates of a point on the surface of a sphere are given by the equations:

$$x = a \cos\phi \cos\theta, \qquad y = a \cos\phi \sin\theta, \qquad z = a \sin\phi.$$

Deduce (3) from these relations and the equation:

$$ds^2 = dx^2 + dy^2 + dz^2.$$

3. Taking as the coordinates of a point on the surface of a cone (ρ, θ), where ρ is the distance from the vertex and θ is the longitude, show that

$$\tan\omega = \frac{d\rho}{\rho\, d\theta \sin\alpha}. \tag{5}$$

4. Obtain the equation and the length of a rhumb line on the cone.

5. The preceding two questions for a cylinder.

Ans. $\tan\omega = \dfrac{dz}{a\,d\theta}, \qquad ds^2 = a^2 d\theta^2 + dz^2,$ where $r = a$ is the equation of the cylinder in cylindrical coordinates.

8. Mercator's Chart. In mapping the earth on a sheet of paper it is not possible to preserve the shapes of the countries and the islands, the lakes and the peninsulas represented. Some distortion is inevitable, and the problem of cartography is to render its disturbing effect as slight as possible. This demand will be met satisfactorily if we can make the angle at which two curves intersect on the earth's surface go over into the same angle on the map. For then a small triangle on the surface of the earth, made by arcs of great circles, will appear in the map as a small curvilinear triangle having the same angles and almost straight sides, and so it will look very similar to the original triangle. What is true of triangles is true of other small figures, and thus we should get a map in which Cuba will look like Cuba and Iceland like Iceland, though the scale for Cuba and the scale for Iceland may be quite different.

A map meeting the above requirement may be made as follows. Regarding the earth as a perfect sphere, construct a cylinder tangent to the earth along the equator. Then the meridians shall go over into the elements of the cylinder and the parallels of latitude into its circular cross-sections as follows: Let P be an arbitrary point on the earth, Q, its image on the cylinder.

(*a*) Q shall have the same longitude, θ, as P.

(*b*) To the latitude ϕ of P shall correspond a distance z of Q from the equator such that the angle ω which an arbitrary curve C through P makes with the parallel of latitude through P, and the angle ω_1 which the image C_1 of C makes with the circular section of the cylinder through Q, shall be the same. Now from (2)

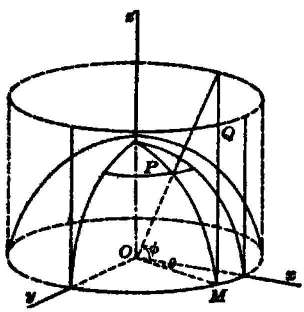


FIG. 41

$$\tan\omega = \frac{d\phi}{d\theta\cos\phi}$$

On the other hand, by § 7, Ex. 5,

$$\tan\omega_1 = \frac{dz}{a\,d\theta}.$$

Hence we get

$$\frac{d\phi}{d\theta\cos\phi} = \frac{dz}{a\,d\theta} \qquad \text{or} \qquad dz = \frac{a\,d\phi}{\cos\phi},$$

$$z = a\int\frac{d\phi}{\cos\phi} = a\log\tan\left(\frac{\phi}{2}+\frac{\pi}{4}\right),$$

the constant of integration vanishing because $z = 0$ corresponds to $\phi = 0$. Thus, if $a = 1$, a point in latitude 60° N. goes over into a point distant 1.32 units from the equator.

The cylinder can now be cut along an element, rolled out on a plane, and the map so obtained reduced to the desired scale.

This map is known as Mercator's Chart.* It has the property that the meridians and the parallels of latitude go over into two orthogonal families of parallel straight lines. Furthermore, a rhumb line on the earth is represented by a straight line on the map. Hence, in order to determine the course of a ship which is to sail from one point to another without altering her course, it is only necessary to lay down a straightedge on the map so that it will go through the two points, and read off the angle it makes with the parallels of latitude.

We call attention to the fact that the above map cannot be obtained by projecting the points of the sphere on the cylinder along a bundle of rays from the centre. It is true that the meridians would go over into a family of parallel straight lines as before, and the same is true of the parallels of latitude, but angles would not in general be preserved.

* G. Kremer, the Latinized form of whose name was Mercator, completed a map of the world on the plan here set forth in 1569.

EXERCISES

1. Turn to an atlas and test the Mercator's charts there found by actual measurement and computation.

2. Show that a curve on the sphere, which cuts the forty-fifth parallel of latitude at an angle of 45°, goes over by the central projection mentioned above into a curve which cuts the image of that parallel of latitude at an angle of 54° 44′.

CHAPTER VII

TAYLOR'S THEOREM. MAXIMA AND MINIMA. LAGRANGE'S MULTIPLIERS

1. The Law of the Mean. Let $f(x, y)$ be a continuous function of the two independent variables x and y, having continuous first partial derivatives throughout a region S. Let (x_0, y_0) be any point within S, and let h, k be two arbitrary constants. Consider

$$f(x_0 + h,\ y_0 + k).$$

We have obtained an expression for this value in terms of $f(x_0, y_0)$ and the first partial derivatives of $f(x, y)$; cf. Chap. V, § 7. The method consisted in forming the function

$$\Phi(t) = f(x_0 + th,\ y_0 + tk), \qquad 0 \leqq t \leqq 1.$$

For,

$$\Phi(1) = f(x_0 + h,\ y_0 + k), \qquad \Phi(0) = f(x_0,\ y_0),$$

and the Law of the Mean for functions of a single variable, applied to $\Phi(t)$, gave:

$$\Phi(1) = \Phi(0) + 1 \cdot \Phi'(\theta), \qquad 0 < \theta < 1.$$

Hence the desired formula resulted, namely:

$$f(x_0 + h,\ y_0 + k) =$$

$$f(x_0, y_0) + hf_1(x_0 + \theta h,\ y_0 + \theta k) + kf_2(x_0 + \theta h,\ y_0 + \theta k),$$

where $0 < \theta < 1$. This is the Law of the Mean for functions of two independent variables. It has been tacitly assumed that the restrictions on the function hold at least throughout the region

$$x = x_0 + th, \qquad y = y_0 + sk, \qquad 0 \leqq t \leqq 1, \qquad 0 \leqq s \leqq 1,$$

s and t being independent of each other.

The extension to functions of $n > 2$ variables is obvious.

2. Taylor's Theorem. We obtain Taylor's Theorem with the Remainder for functions of several variables if we write the corresponding theorem for $\Phi(t)$, *Introduction to the Calculus*, p. 430:

$$\Phi(1) = \Phi(0) + \Phi'(0) + \cdots + \frac{1}{n!}\Phi^{(n)}(0) + \frac{1}{(n+1)!}\Phi^{(n+1)}(\theta),$$

and then substitute for Φ and its derivatives their values. Thus when $n = 1$ we get

$$(1) \qquad f(x_0+h, y_0+k) = f(x_0, y_0) + hf_1(x_0, y_0) + kf_2(x_0, y_0)$$
$$+ \tfrac{1}{2}[h^2 f_{11}(X, Y) + 2hk f_{12}(X, Y) + k^2 f_{22}(X, Y)],$$

where $X = x_0 + \theta h$, $Y = y_0 + \theta k$, and $0 < \theta < 1$.

The student should write out the formula for the next case, $n = 2$.

The general term, $\Phi^{(n)}(0)/n!$, can be expressed symbolically as

$$\frac{1}{n!}\left[h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right]^n f(x, y)\Big|_{\substack{x=x_0\\ y=y_0}}$$

and the remainder as

$$\frac{1}{(n+1)!}\left[h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right]^{n+1} f(x, y)\Big|_{\substack{x=x_0+\theta h\\ y=y_0+\theta k}}$$

The extension to functions of $n > 2$ variables is immediate.

If the remainder converges toward zero when n becomes infinite, we obtain an infinite series whose terms are homogeneous polynomials and which converges toward the value of the function. If, furthermore, the series whose terms consist of the monomials that make up the terms of the latter series converges for all values of h and k within certain limits: $|h| < H$, $|k| < K$, we say that the function can be *developed into a power series in* $h = x - x_0$ *and* $k = y - y_0$:

$$(2) \qquad f(x, y) = \sum c_{mn}(x - x_0)^m (y - y_0)^n,$$

or that it can be *developed by Taylor's Theorem.* A series of the form (2) is often called a Taylor's Series. But it is not in general feasible to give a direct proof that the remainder converges toward zero, and so other methods of analysis have to be employed to establish a Taylor's development.

EXERCISE

Assuming that the function $e^x \cos y$ can be developed by Taylor's Theorem about the point $x_0 = y_0 = 0$, show that

$$e^x \cos y = 1 + x + \tfrac{1}{2}(x^2 - y^2) + \cdots.$$

3. Maxima and Minima. The function $f(x, y)$ will have a maximum at the point (x_0, y_0) if the tangent plane of the surface

$$u = f(x, y)$$

at (x_0, y_0) is parallel to the (x, y)-plane and the surface lies below this plane at all other points of the neighborhood of (x_0, y_0, u_0). Hence we see that at (x_0, y_0)

$$(1) \qquad \frac{\partial u}{\partial x} = 0, \qquad \frac{\partial u}{\partial y} = 0.$$

A similar statement holds for a minimum.*

The necessary condition contained in (1) can be extended at once to functions of $n > 2$ variables. For, if any one of the first partial derivatives, as $\partial u/\partial x$, for example, were $\neq 0$ at $(x_0, y_0, z_0, \cdots)$, then the function $f(x, y_0, z_0, \cdots)$, which is a function of x alone, would be increasing as x passes through the value x_0, or else it would be decreasing, according to the sign of $\partial u/\partial x$, and so in neither case could $f(x, y, z, \cdots)$ have a maximum or a minimum there.

The conditions (1) are frequently sufficient, together with other information, to determine a maximum or a minimum.

Example 1. Given three particles of masses m_1, m_2, m_3, situated at the points (x_1, y_1), (x_2, y_2), (x_3, y_3). To find the point about which the moment of inertia of these particles will be a minimum.

Here it is clear that for all distant points of the plane the moment of inertia is large, becoming infinite in the infinite region of the plane. Furthermore, the moment of inertia, I, is a positive continuous function. Hence the surface $u = I$, or

$$u = m_1[(x - x_1)^2 + (y - y_1)^2] + m_2[(x - x_2)^2 + (y - y_2)^2] + m_3[(x - x_3)^2 + (y - y_3)^2]$$

must have at least one minimum, and at such a point

$$\frac{\partial u}{\partial x} = 2[m_1(x - x_1) + m_2(x - x_2) + m_3(x - x_3)] = 0,$$

$$\frac{\partial u}{\partial y} = 2[m_1(y - y_1) + m_2(y - y_2) + m_3(y - y_3)] = 0.$$

But these equations determine the centre of gravity of the particles and are satisfied by no other point. Hence the centre of gravity is the point about which the moment of inertia is least.

The result is in accordance with the general theorem of the *Introduction to the Calculus*, p. 331, and it holds for any system of particles whatever.

* A point (x_0, y_0) at which equations (1) alone are satisfied — apart from any further condition — may be called a point at which the function is *stationary*, since the change in value which the function experiences when (x, y) moves to a point $(x_0 + h, y_0 + k)$ near by is an infinitesimal of higher order than the distance between these points.

Auxiliary Variables. As in the case of functions of a single variable, so here it frequently happens that it is best to express the quantity to be made a maximum or a minimum in terms of more variables than are necessary, one or more relations existing between these variables. The student must, therefore, in all cases begin by considering *how many independent variables* there are, and then write down *all* the relations between the letters that enter; and he must make up his mind as to what letters he will take as independent variables before he begins to differentiate.

Example 2. What is the volume of the greatest rectangular parallelepiped that can be inscribed in the ellipsoid:

(2) $$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1\,?$$

We assume that the faces are to be parallel to the coordinate planes and thus obtain for the volume:

$$V=8\,xyz.$$

But x, y, z cannot all be chosen at pleasure. They are connected by the relation (2). So the number of independent variables is here two, and we may take them as x and y. We have, then:

(3) $$\frac{\partial V}{\partial x}=8\,y\left(z+x\frac{\partial z}{\partial x}\right)=0, \qquad \frac{\partial V}{\partial y}=8\,x\left(z+y\frac{\partial z}{\partial y}\right)=0.$$

From (2) we obtain:

(4) $$\frac{\partial z}{\partial x}=-\frac{c^2x}{a^2z}, \qquad \frac{\partial z}{\partial y}=-\frac{c^2y}{b^2z}.$$

Now, neither $x=0$ nor $y=0$ can lead to a solution of the problem, and hence it follows from (3) and (4) that

$$z-\frac{c^2x^2}{a^2z}=0, \qquad z-\frac{c^2y^2}{b^2z}=0,$$

or

$$\frac{x^2}{a^2}=\frac{y^2}{b^2}=\frac{z^2}{c^2}.$$

Thus the parallelepiped whose vertices lie at the intersections of these lines with the ellipsoid, *i.e.* on the diagonals of the circumscribed parallelepiped $x=\pm a$, $y=\pm b$, $z=\pm c$, is the one required,* and its volume is

$$V=\tfrac{8}{9}\sqrt{3}\,abc.$$

* The reasoning, given at length, is as follows. V is a continuous positive function of x and y at all points within the ellipse

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1,$$

EXERCISES

1. Required the parallelepiped of given volume and minimum surface. *Ans.* A cube.

2. Required the parallelepiped of given surface and maximum volume. *Ans.* A cube.

3. A tank in the form of a rectangular parallelepiped, open at the top, is to be built, and it is to hold a given amount of water. Find what proportions it should have, in order that the cost of lining it may be as small as possible. How many independent variables are there in this problem?

Ans. Length and breadth each double the depth.

4. Find the shortest distance between the lines

$$\begin{cases} y = 2x, \\ z = 5x, \end{cases} \qquad \begin{cases} y = 3x + 7 \\ z = x. \end{cases}$$

5. Show without using the calculus that the function

$$x^4 + y^4 + 4x - 32y - 7$$

has a minimum.

Suggestion. Use polar coordinates.

6. Find the minimum in the preceding problem.

7. A hundred tenement houses of given cubical content are to be built in a factory town. They are to have a rectangular ground plan and a gable roof. Find the dimensions for which the area of walls and roof will be least.*

8. A torpedo in the form of a cylinder with equal conical ends is to be made out of boiler plates and is just to float when loaded. The displacement of the torpedo being given, what must be its proportions, that it may carry the greatest weight of dynamite?

Ans. The length of the torpedo must be three times the length of the cylindrical portion, and the diameter must be $\sqrt{5}$ times the length of the cylindrical portion.

---

for which $x > 0$, $y > 0$, and it vanishes on the boundary of this region. Hence it must have at least one maximum inside. But we find only one point, namely $x = a/\sqrt{3}$, $y = b/\sqrt{3}$, at which V can possibly be at a maximum. Hence, etc.

* The problem is identical with that of finding the best shape for a wall-tent.

9. Find the point so situated that the sum of its distances from the three vertices of an acute-angled triangle is a minimum.

Ans. The lines joining the point with the vertices make angles of 120° with one another.*

10. Find the most economical dimensions for a powder house of given cubical content, if it is built in the form of a cylinder and the roof is a cone.

4. Test by the Derivatives of the Second Order. We proceed to deduce a sufficient condition for a relative maximum or minimum in terms of the derivatives of the second order. Suppose the necessary conditions, § 3, (1) are fulfilled at (x_0, y_0). Then from § 2, (1) we get:

(1) $$f(x_0+h, y_0+k)-f(x_0, y_0)=\tfrac{1}{2}(Ah^2+2Bhk+Ck^2),$$

where $A=f_{11}(x_0+\theta h, y_0+\theta k)$, $B=f_{12}(x_0+\theta h, y_0+\theta k)$,

$$C=f_{22}(x_0+\theta h, y_0+\theta k),$$

and for a minimum the difference (1) must be positive for all points $x=x_0+h$, $y=y_0+k$ near (x_0, y_0) except for this one point, where it vanishes.

Definite Quadratic Forms. A homogeneous polynomial of the second degree in any number of variables is called a *quadratic form*,† and is said to be *definite* if it vanishes only when all the variables vanish; otherwise it is said to be *indefinite*. Thus

$$h^2+k^2, \qquad 2h^2+3k^2+5l^2$$

are examples of definite quadratic forms in two and three variables respectively;

$$h^2, \qquad 3h^2+7hk+2k^2=(3h+k)(h+2k)$$

are indefinite. A definite quadratic form never changes sign; an indefinite one may.

THEOREM. *In order that*

$$U=Ah^2+2Bhk+Ck^2,$$

* For a complete discussion of the problem for any triangle see Goursat-Hedrick, *Mathematical Analysis*, vol. 1, § 62.

† For some purposes it is desirable to define an *algebraic form* merely as a polynomial. But we are concerned here only with homogeneous polynomials. Moreover, we exclude the case that all the coefficients vanish.

where A, B, C are independent of h and k, be a definite form, it is necessary and sufficient that

$$(2) \qquad B^2 - AC < 0.$$

That this condition is sufficient is at once evident. For, if it is fulfilled, surely neither A nor C can vanish, and we can write:

$$U = \frac{1}{A}[(Ah + Bk)^2 + (AC - B^2)k^2].$$

Hence U can vanish only when

$$Ah + Bk = 0 \qquad \text{and} \qquad k = 0,$$

i.e. only when $h = k = 0$, q. e. d.

We leave the proof that the condition is necessary to the student.

When the condition (2) is fulfilled, A and C necessarily have the same sign, and this is the sign of U.

COROLLARY. *If A, B, C depend on h and k in any manner whatever, and if, for a pair of values (h, k) not both zero, the condition (2) is fulfilled, then for these values U has the same sign as A and C.*

Application to Maxima and Minima. Returning now to equations (1), let us suppose that

$$(3) \qquad \left(\frac{\partial^2 u}{\partial x\,\partial y}\right)^2 - \frac{\partial^2 u}{\partial x^2}\frac{\partial^2 u}{\partial y^2} < 0$$

at (x_0, y_0) and that these derivatives are continuous in the vicinity of this point. Then the relation (3) will hold for all points near (x_0, y_0) and furthermore, for such points, both $\frac{\partial^2 u}{\partial x^2}$ and $\frac{\partial^2 u}{\partial y^2}$ will preserve the common sign they have at (x_0, y_0). Hence the right-hand side of (1) will vanish only at (x_0, y_0), and at other points in the neighborhood will have the sign of these latter derivatives. We are thus led to the following

SUFFICIENT CONDITION FOR A RELATIVE MAXIMUM OR MINIMUM. *If at the point (x_0, y_0)*

$$(a) \qquad \frac{\partial u}{\partial x} = 0, \qquad \frac{\partial u}{\partial y} = 0;$$

$$(b) \qquad \left(\frac{\partial^2 u}{\partial x\,\partial y}\right)^2 - \frac{\partial^2 u}{\partial x^2}\frac{\partial^2 u}{\partial y^2} < 0;$$

and if the derivatives of the second order are continuous near (x_0, y_0), then u will have a relative maximum at (x_0, y_0) if

$$(c_1) \qquad \frac{\partial^2 u}{\partial x^2} < 0,$$

and a relative minimum there if

$$(c_2) \qquad \frac{\partial^2 u}{\partial x^2} > 0.$$

Conditions (b) and (c) are not necessary, but only sufficient. u may have a maximum or a minimum even when the sign of inequality in (b) is replaced by the sign of equality. But if, in (b), the sign of inequality be reversed, u has neither a maximum nor a minimum.

When f depends on $n > 2$ variables, the method of procedure is similar. First, the algebraic theorem about quadratic forms has to be generalized. Thus for three variables,

$$(4) \qquad U = a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + 2\,a_{23}x_2x_3 + 2\,a_{31}x_3x_1 + 2\,a_{12}x_1x_2$$

and a necessary and sufficient condition that U be a positive definite quadratic form is that

$$(5) \qquad a_{11} > 0, \qquad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \qquad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} > 0,$$

where $a_{ij} = a_{ji}$. This form of statement suggests the generalization for $n = n$.

If U is to be a negative definite quadratic form, the first, third, fifth, etc. inequality signs in (5) must be reversed.

EXERCISES

1. Show that the surface

$$z = xy$$

has neither a maximum nor a minimum at the origin.

2. Show that the function

$$x^3 + 3\,x^2 - 2\,xy + 5\,y^2 - 4\,y^3$$

has a relative minimum at the origin.

3. Test the function

$$2\,x^2 + 2\,xy + 5\,y^2 + 2\,x - 2\,y + 1$$

for maxima and minima.

4. Determine the maxima and minima of the surface

$$x^2 + 2\,y^2 + 3\,z^2 - 2\,xy - 2\,yz = 2.$$

5. Lagrange's Multipliers. Let it be required to find the condition that the function

(1) $$u = F(x, y, z)$$

may be stationary, where x, y, z are connected by a relation

(2) $$\Phi(x, y, z) = 0.$$

The condition is represented by equations (1) of § 3, extended to this more general case.

If equation (2) can be solved for z, then, on substituting this value for z in (1), we have u expressed as a function of the two independent variables x and y, and hence equations (1) of § 3 become

(3) $$\frac{\partial u}{\partial x} = F_1 + F_3\frac{\partial z}{\partial x} = 0, \qquad \frac{\partial u}{\partial y} = F_2 + F_3\frac{\partial z}{\partial y} = 0.$$

On the other hand, we have for the determination of $\partial z/\partial x$ and $\partial z/\partial y$ from (2) the equations

(4) $$\Phi_1 + \Phi_3\frac{\partial z}{\partial x} = 0, \qquad \Phi_2 + \Phi_3\frac{\partial z}{\partial y} = 0.$$

From the pair of equations consisting of the first equation in (3) and (4) $\partial z/\partial x$ can be eliminated:

(5) $$\begin{vmatrix} F_1 & F_3 \\ \Phi_1 & \Phi_3 \end{vmatrix} = 0;$$

and similarly, from the pair consisting of the second equation, $\partial z/\partial y$ can be eliminated:

(6) $$\begin{vmatrix} F_2 & F_3 \\ \Phi_2 & \Phi_3 \end{vmatrix} = 0.$$

Equations (5) and (6), combined with (2), are three equations for determining the values of the three unknowns x, y, and z, for which equations (3) and (4) hold simultaneously, and hence for which equations (1) of § 3 are true.

Lagrange observed that the problem just solved is equivalent to the following problem. Let u be set equal to the function $F + \lambda\Phi$:

(7) $$u = F(x, y, z) + \lambda\Phi(x, y, z),$$

where λ is a constant, to which we will later assign a suitable value. Consider u as a function now of the *three* independent variables, x, y, and z,* and write down the conditions corresponding to (1) of

* This step is wholly arbitrary. The *motif* lies in the purely algebraic situation which arises when we do this thing.

§ 3 for this function; thus we have

$$(8) \qquad F_1 + \lambda\Phi_1 = 0, \qquad F_2 + \lambda\Phi_2 = 0, \qquad F_3 + \lambda\Phi_3 = 0.$$

If these equations are to hold simultaneously, then it follows from the third of them that λ must have the value:

$$(9) \qquad \lambda = -\frac{F_3}{\Phi_3}.$$

Hence, on substituting this value of λ in each of the first two of the equations (8) we are led to equations (5) and (6).

Thus we obtain by Lagrange's method these two pivotal equations, and no more; for equation (9) imposes no condition on x, y, z, but serves merely to determine λ. Coming back, now, to the original problem of making u as given by (1) under the restriction (2) a maximum or a minimum, we add to the two equations just mentioned equation (2), and thus have the same system of three equations which we obtained by the first method.

The method can be extended at once to the case of a function u of n variables, $x_1, \cdots, x_n$, which are connected by a single relation:

$$(10) \qquad u = F(x_1, \cdots, x_n), \qquad \Phi(x_1, \cdots, x_n) = 0.$$

The last equation can equally well be written in the form

$$\Phi(x_1, \cdots, x_n) = C,$$

where C is any constant. — Cf. further § 6.

Example. (See also p. 526.) Let

$$F(x, y, z) = Ax^2 + By^2 + Cz^2 + 2\,Dyz + 2\,Ezx + 2\,Fxy$$

be a positive definite quadratic form, and let x, y, z be restricted by the relation

$$\Phi(x, y, z) = x^2 + y^2 + z^2 = a^2.$$

Equations (8) now take on the form

$$(11) \qquad \begin{cases} Ax + Fy + Ez + \lambda x = 0, \\ Fx + By + Dz + \lambda y = 0, \\ Ex + Dy + Cz + \lambda z = 0. \end{cases}$$

A necessary condition that these equations hold simultaneously is that

$$(12) \qquad \begin{vmatrix} A+\lambda & F & E \\ F & B+\lambda & D \\ E & D & C+\lambda \end{vmatrix} = 0.$$

If, in particular,

(13) $$F(x, y, z) = Ax^2 + By^2 + Cz^2,$$

the equation (12) reduces to

$$(A+\lambda)(B+\lambda)(C+\lambda) = 0,$$

and the roots, which are all real, are $-A$, $-B$, and $-C$.

In the general case, equation (12) has at least one real root, and no root of (12) is 0. For, the vanishing of the determinant arising by setting $\lambda = 0$ in (12) would mean that $F(x, y, z)$ could vanish for values of x, y, z not all 0.

Consider the function $F(x, y, z)$ in the points of the sphere $\Phi = a^2$. Since $F(x, y, z)$ is continuous, it has a maximum value on the sphere, and also a minimum value.* Hence two of the three roots of (12) are real and distinct, and thus all three are real.

Let the coordinate axes be so rotated that F attains its maximum value in the point $(0, 0, \zeta)$, where $\zeta > 0$. From (11) it appears that $D = 0$, $E = 0$, $\lambda = -C$. Thus

$$F(x, y, z) = Ax^2 + 2Fxy + By^2 + Cz^2, \qquad C > 0.$$

If $F \neq 0$, a suitable rotation of the axes about the axis of z will remove the term in xy, and thus the form (13) is attained, where A, B, C are all positive.

EXERCISES

1. Find the values of (x, y, z) for which the function

$$u = xyz$$

is stationary, if $$x + y + z = 1.$$

2. Work Example 2 of § 3 by means of Lagrange's multipliers.

3. Examine the Exercises at the close of § 3 and determine to which of these the method of Lagrange's multipliers is particularly well adapted.

4. Show that the method of Lagrange's multipliers is valid in the case of functions of a single variable, given in the form:

$$u = F(x, y) \qquad \Phi(x, y) = 0.$$

* If, in particular, these values are the same, *i.e.* if $F(x, y, z)$ is constant on the sphere, then $F(x, y, z) \equiv K\Phi(x, y, z)$ ($K =$ const.) and all three roots of (12) are real and equal.

The condition here takes on the form:

$$\begin{vmatrix} F_1 & F_2 \\ \Phi_1 & \Phi_2 \end{vmatrix} = 0.$$

5. Show that, if

$$u = F(x_1, \cdots, x_n) = \sum a_{ij} x_i x_j \quad \text{and} \quad \Phi(x_1, \cdots, x_n) = x_1^2 + \cdots + x_n^2 = a^2,$$

where i and j, independently of each other, run through the values from 1 to n, and $a_{ij} = a_{ji}$, equations (8) assume the form

$$\text{(14)} \qquad \begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n + \lambda x_1 = 0, \\ a_{21}x_1 + \cdots + a_{2n}x_n + \lambda x_2 = 0, \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot, \\ a_{n1}x_1 + \cdots + a_{nn}x_n + \lambda x_n = 0. \end{cases}$$

Hence λ must be a root of the equation

$$\text{(15)} \qquad \begin{vmatrix} a_{11} + \lambda & a_{12} & \cdot & \cdot & \cdot & \cdot & a_{1n} \\ a_{21} & a_{22} + \lambda & \cdot & \cdot & \cdot & \cdot & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \cdot & \cdot & \cdot & \cdot & a_{nn} + \lambda \end{vmatrix} = 0.$$

6. Show that if, in the preceding question, the function u has a maximum or a minimum in the point $(0, 0, \cdots, 0, x_n')$, where $x_n' > 0$, then F contains only a simple term involving x_n, namely $a_{nn}x_n^2$.

7. A "rotation" of space of m dimensions is given by the formula

$$x_k' = a_{k1}x_1 + \cdots + a_{km}x_m, \qquad k = 1, \cdots, m,$$

where

$$\sum_{l=1}^{m} a_{kl}a_{lp} = \begin{cases} 0, & k \neq p \\ 1, & k = p \end{cases} \quad \text{and} \quad a_{kl} = a_{lk},$$

provided the determinant of the transformation (which $= \pm 1$) has the value $+1$.

Show that, by a succession of rotations (which can, of course, be compounded into a single rotation) the form F of Question 5 can be carried into a form in which only the terms in x_k^2 are present.

8. Find the points of the circle

$$x^2 + y^2 + z^2 = 1, \qquad ax + by + cz = 0,$$

in which the function

$$w = Ax^2 + By^2 + Cz^2 + 2Dyz + 2Ezx + 2Fxy$$

attains its greatest and its least values. Treat first the case:

$$D = E = F = 0.$$

6. Continuation. Several Auxiliary Equations. The method of Lagrange's multipliers applies to the general case that the variables are connected by an arbitrary number of auxiliary equations. For example, let

$$(1) \qquad u = F(x, y, z, t),$$

$$(2) \qquad \Phi(x, y, z, t) = 0, \qquad \Psi(x, y, z, t) = 0.$$

If equations (2) can be solved for z and t, then, on substituting these values in (1), u becomes a function of x and y. Equations (1) of § 3 now take on the form

$$(3) \qquad F_1 + F_3\frac{\partial z}{\partial x} + F_4\frac{\partial t}{\partial x} = 0, \qquad F_2 + F_3\frac{\partial z}{\partial y} + F_4\frac{\partial t}{\partial y} = 0.$$

The derivatives $\partial z/\partial x$, etc. are determined by the equations:

$$(4) \qquad \Phi_1 + \Phi_3\frac{\partial z}{\partial x} + \Phi_4\frac{\partial t}{\partial x} = 0, \qquad \Phi_2 + \Phi_3\frac{\partial z}{\partial y} + \Phi_4\frac{\partial t}{\partial y} = 0;$$

$$(5) \qquad \Psi_1 + \Psi_3\frac{\partial z}{\partial x} + \Psi_4\frac{\partial t}{\partial x} = 0, \qquad \Psi_2 + \Psi_3\frac{\partial z}{\partial y} + \Psi_4\frac{\partial t}{\partial y} = 0.$$

Thus we find the conditions

$$(6) \qquad \begin{vmatrix} F_1 & F_3 & F_4 \\ \Phi_1 & \Phi_3 & \Phi_4 \\ \Psi_1 & \Psi_3 & \Psi_4 \end{vmatrix} = 0, \qquad \begin{vmatrix} F_2 & F_3 & F_4 \\ \Phi_2 & \Phi_3 & \Phi_4 \\ \Psi_2 & \Psi_3 & \Psi_4 \end{vmatrix} = 0.$$

The four equations (2) and (6) determine the four unknowns x, y, z, t, and for this system of values equations (1) of § 3 hold.

Lagrange's method consists in forming the function

$$u = F + \lambda\Phi + \mu\Psi,$$

where λ and μ are constants, to which shall later be assigned suitable values, and where u is considered as a function of the *four* independent variables, (x, y, z, t). It is to this function that condition (1) of § 3 is now applied, and thereby result the equations:

$$(7) \quad F_1 + \lambda\Phi_1 + \mu\Psi_1 = 0, \quad F_2 + \lambda\Phi_2 + \mu\Psi_2 = 0, \quad F_3 + \lambda\Phi_3 + \mu\Psi_3 = 0,$$
$$F_4 + \lambda\Phi_4 + \mu\Psi_4 = 0.$$

From the last two of these λ and μ are to be determined, and these values are then substituted in the first two. The two equations thus obtained are precisely the equations (6).

The extension of the method to a function of n variables,

$$u = F(x_1, \cdots, x_n),$$

the variables being connected by p equations:

$$\Phi^{(1)}(x_1, \cdots, x_n) = 0, \quad \cdot \quad \cdot \quad \cdot \quad , \quad \Phi^{(p)}(x_1, \cdots, x_n) = 0,$$

is now obvious. The relations which correspond to (6) are:

$$(8) \qquad \begin{vmatrix} F_k & F_{n-p+1} \cdots F_n \\ \Phi_k^{(1)} & \Phi_{n-p+1}^{(1)} \cdots \Phi_n^{(1)} \\ \cdot & \cdot \quad \cdot \quad \cdot \quad \cdot \\ \Phi_k^{(p)} & \Phi_{n-p+1}^{(p)} \cdots \Phi_n^{(p)} \end{vmatrix} = 0, \qquad k = 1, \cdots, p.$$

7. Conclusion; Critique. In the first case considered, § 5, it was tacitly assumed that the functions $F(x, y, z)$ and $\Phi(x, y, z)$ are continuous, together with their first partial derivatives, in the neighborhood of a point (x_0, y_0, z_0) whose coordinates satisfy equations (2), (5), and (6). But this is not enough. The equation (2) must determine such a function z of x and y that equations (3) can have a meaning. This will surely be the case if $\Phi_3(x_0, y_0, z_0) \neq 0$. Moreover, this is also precisely the condition which we need in Lagrange's method, in order that equation (9) may have a meaning. It is, of course, immaterial whether we solve equation (2) for z or for one of the other letters. We see, then, that Lagrange's method will apply *if at least one of the numbers* $\Phi_k(x_0, y_0, z_0)$, $k = 1, 2, 3$, *is different from* 0.

In § 6 the situation is similar. It is enough, in addition to the continuity of the functions F, Φ, Ψ (together with that of their first partial derivatives) in the neighhorhood of a point whose coordinates satisfy equations (2) and (6), that at least one of the two-rowed determinants

$$\begin{vmatrix} \Phi_i & \Phi_j \\ \Psi_i & \Psi_j \end{vmatrix},$$

where i and j are two distinct numbers chosen from the set 1, 2, 3, be different from zero.

The extension to the general case is now obvious. At least one p-rowed determinant from the matrix made up of the last p rows of the determinant (8) must be different from zero; — at least, this is sufficient, in order that u be stationary. The student must have a firm hold on the theory of Linear Dependence; cf. Bôcher, *Algebra*, Chaps. 3, 4.

CHAPTER VIII

ENVELOPES

1. Envelope of a Family of Curves. Consider a family of circles, of equal radii, whose centres all lie on a right line. The family is represented by the equation

(1) $$(x-\alpha)^2+y^2=1.$$

where the parameter α runs through all values. The lines

(2) $$y=1 \qquad \text{and} \qquad y=-1$$

are touched by all the curves of this family.

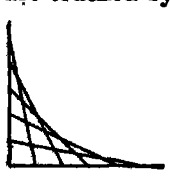


FIG. 42

Again, let a rod slide with one end on the floor and the other touching a vertical wall, the rod always remaining in the same vertical plane. It is clear that the rod in its successive positions is always tangent to a certain curve. This curve, like the lines (2) in the preceding example, is called the *envelope* of the family of curves.

Turning now to the general case, we see that the family of curves

(3) $$f(x, y, \alpha)=0$$

may have one or more curves to which, as α varies, the successive members of the family are tangent. When this is so, two curves of the family corresponding to values of α differing but slightly from each other:

(4) $$f(x, y, \alpha_0)=0, \qquad f(x, y, \alpha_0+\Delta\alpha)=0,$$

will usually intersect near the points of contact of these curves with the envelope, as is illustrated in the above examples. So if we determine the limiting position of this point P of intersection of the curves (4), we shall obtain a point of the envelope. We will first outline the method and show its application, and then come back to a study of the details in § 4.

From analytic geometry * we know that, if $u=0$ and $v=0$ are

* *Analytic Geometry*, p. 165.

the equations of two curves, then $u + kv = 0$ (where k is a constant) represents a curve which passes through all the points of intersection of the given curves. Applying this principle to the curves (4), we see that a third curve through P is given by the equation

(5) $$f(x, y, \alpha_0 + \Delta\alpha) - f(x, y, \alpha_0) = 0.$$

The left-hand side has the value $\Delta\alpha f_\alpha(x, y, \alpha_0 + \theta\Delta\alpha)$ (Law of the Mean, Chap. V, § 2). Hence the coordinates of P satisfy the equation

(6) $$f_\alpha(x, y, \alpha_0 + \theta\Delta\alpha) = 0.$$

Now let $\Delta\alpha$ approach 0 as its limit. The point P approaches the point of tangency of the first curve (4) with the envelope, and the left-hand side of (6) approaches $f_\alpha(x, y, \alpha_0)$. Hence the equation

$$f_\alpha(x, y, \alpha_0) = 0$$

represents a second curve passing through the point of tangency of the first curve (4) with the envelope. Thus we obtain the

THEOREM. *The envelope of the family of curves*

$$f(x, y, \alpha) = 0$$

is given by the pair of equations

(7) $$f(x, y, \alpha) = 0, \qquad \frac{\partial f}{\partial \alpha} \equiv f_\alpha(x, y, \alpha) = 0.$$

Example 1. Applying formulas (7) to the family of circles (1) we get:

$$\frac{\partial f}{\partial \alpha} = -2(x - \alpha) = 0.$$

Equations (7) now tell us that the envelope is given by the pair of equations

$$(x - \alpha)^2 + y^2 = 1, \qquad x - \alpha = 0.$$

These equations are equivalent to the single equation obtained by eliminating α:

$$y^2 = 1, \qquad \text{or} \qquad y = 1 \quad \text{and} \quad y = -1.$$

The analytic result is seen to correspond with the geometric evidence.

Example 2. To find the envelope of the family of ellipses whose axes coincide and whose areas are constant.

Here,

(a) $$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1,$$

(b) $$\pi ab = k.$$

It is more convenient to retain both parameters, rather than to eliminate, but we must be careful to remember that only one is independent. If we choose a as that one, $a = \alpha$, and differentiate with respect to α, we have:

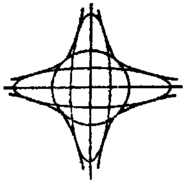


FIG. 43

$$-\frac{2x^2}{a^3}-\frac{2y^2}{b^3}\frac{db}{da}=0, \qquad \pi\left(b+a\frac{db}{da}\right)=0,$$

and hence

(c) $$\frac{x^2}{a^2}=\frac{y^2}{b^2}.$$

Between (a), (b), and (c) we can eliminate a and b and thus get a single equation in x and y, which will be the equation of the envelope. To do this, solve (a) and (c) for a^2 and b^2, thus getting

$$a^2 = 2x^2, \qquad b^2 = 2y^2;$$

and then substitute the values of a and b from these equations in (b):

$$\pm 2\pi xy = k,$$

This equation represents a pair of equal equilateral hyperbolas on the axes as asymptotes.

The equations

$$x = \pm a/\sqrt{2}, \qquad y = \pm b/\sqrt{2},$$

combined with (b), give the coordinates of the points of the envelope in which the particular ellipse corresponding to that pair of values of a and b is tangent to it. This remark applies generally whenever the coordinates x and y of a point of the envelope are obtained as functions of α.

EXERCISES

1. Find the envelope of the family of straight lines

$$2\alpha y = 2x + \alpha^2.$$

Draw a number of the lines.

2. The same question for the family

$$x\cos\alpha + y\sin\alpha = 2.$$

3. The legs of a right triangle lie along two fixed lines, and the

hypotenuse varies so that the area of the triangle is always the same. Find the envelope of the hypotenuse.

Draw an accurate figure showing a good number of the triangles. Take 1 cm. as the unit of length and 2 sq. cm. as the area of the triangles.

4. Circles are drawn on the chords of a parabola which are perpendicular to the axis, as diameters. Show that the envelope is an equal parabola.

Make an accurate drawing of a good number of circles for the parabola $y^2 = x$, 1 cm. being taken as the unit of length.

5. Show that the envelope of all ellipses having coincident axes, the distance between two consecutive vertices of any ellipse being the same for all the ellipses, is a square.

6. What is the envelope of all the chords of a circle which are of a given length?

7. Find the envelope of straight lines drawn perpendicular to the normals of a parabola at the points where they cut the axis.

8. Find the envelope of a circle which is always tangent to the axis of x and always has its centre on the parabola $y = x^2$.

9. Show that the envelope of the lines in the second example of § 1 is an arc of a four-cusped hypocycloid.

10. A straight line moves in such a way that the sum of its intercepts on two rectangular axes is constant. Find its envelope. Draw an accurate figure.

Observe that the equation $\sqrt{x} + \sqrt{y} = \sqrt{l}$ represents a parabola.

11. Find the envelope of the family of circles which pass through the origin and have their centres on the hyperbola $xy = 1$.

Ans. The lemniscate $(x^2 + y^2)^2 = 16\,xy$.

12. The streams of water in a fountain issue from the nozzle, which is small, in all directions, but with the same velocity, v_0. Show that, if the nozzle be regarded as a point, the form of the fountain is a paraboloid of revolution.

13. Circles are drawn on those chords of an ellipse as diameter which are parallel to an axis of the ellipse. Show that the envelope is part of an ellipse, one axis of which is equal to the axis mentioned, the other axis being equal to the diagonal of the rectangle which circumscribes the ellipse.

14. Circles are drawn on chords of the hyperbola $xy = 1$ as diameters. Find their envelope.

2. Envelope of Tangents and Normals. Any curve may be regarded as the envelope of its tangents. Thus the equation of the tangent to the parabola

(1) $$y^2 = 2mx$$

at the point (x_0, y_0) is

$$y - y_0 = \frac{m}{y_0}(x - x_0)$$

or

(2) $$y = \frac{mx}{y_0} + \frac{y_0}{2}.$$

Hence the envelope of the lines (2), where y_0 is regarded as a parameter, must be the parabola (1), and the student can readily assure himself that this is the case.

The evolute of a curve was defined as the locus of the centres of curvature, and it was shown that the normal to the curve is tangent to the evolute; *Introduction to the Calculus*, p. 266, § 4. Hence the evolute is the envelope of the normals, and thus we have a new method for determining the evolute.

For example, the equation of the normal to the parabola

$$y = x^2$$

at the point (x_0, y_0) is

$$x - x_0 + 2x_0(y - y_0) = 0$$

or

$$x + 2x_0 y - x_0 - 2x_0^3 = 0,$$

and we get at once as the envelope of this family of lines:

$$y = 3x_0^2 + \tfrac{1}{2}, \qquad x = -4x_0^3,$$

or

$$(y - \tfrac{1}{2})^3 = \tfrac{27}{16}x^2.$$

The result agrees with that obtained, *l.c.*, p. 264.

EXERCISES

1. Obtain the equation of the evolute of the ellipse:

$$x = a\cos\phi, \qquad y = b\sin\phi,$$

as the envelope of its normals.

2. The same question for the hyperbola

$$x = a\sec\phi, \qquad y = b\tan\phi.$$

3. Obtain the evolute of the cycloid:

$$x = a(\theta - \sin\theta), \qquad y = a(1 - \cos\theta).$$

4. Obtain the coordinates (x_1, y_1) of any point on the envelope of the normals to the curve $y = f(x)$:

$$x - x_0 + f'(x_0)(y - y_0) = 0,$$

and show that the result agrees with the formulas of the *Introduction to the Calculus*, p. 263, (9).

5. Obtain the evolute of the family of lines

$$(b^2 - a^2\lambda^2)\,x + 2\,a^2\lambda y = a(b^2 + a^2\lambda^2).$$

Ans. The ellipse, $x = \dfrac{a(b^2 - a^2\lambda^2)}{b^2 + a^2\lambda^2}, \qquad y = \dfrac{2\,ab^2\lambda}{b^2 + a^2\lambda^2}.$

3. Caustics. When rays of light that are nearly parallel fall on the concave side of a napkin ring or a water glass, a portion of the table cloth is illuminated. Let us determine the equation of the boundary.

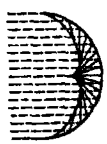


FIG. 44

Suppose we have a narrow semicircular band, on the polished concave side of which a bundle of parallel rays fall. The rays are reflected at the same angle with the normal as the angle of incidence, and so we wish to find the envelope of the reflected rays. Take the radius of the band as 1. Then the equation of the reflected ray is

(1) $$y - \sin\theta = \tan 2\theta\,(x - \cos\theta).$$

To get the envelope of the family, we differentiate with respect to θ:

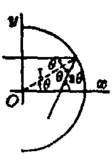


FIG. 45

$$-\cos\theta = 2\sec^2 2\theta\,(x - \cos\theta) + \tan 2\theta\,\sin\theta,$$

$$2\,x = 2\cos\theta - \cos^2 2\theta\cos\theta - \cos 2\theta\sin 2\theta\sin\theta$$

$$= 2\cos\theta - \cos 2\theta\,(\cos 2\theta\cos\theta + \sin 2\theta\sin\theta)$$

$$= 2\cos\theta - \cos 2\theta\cos\theta,$$

or: $$x = \tfrac{3}{4}\cos\theta - \tfrac{1}{4}\cos 3\theta.$$

Substituting this value of x in (1) we get:

$$y = \tfrac{3}{4}\sin\theta - \tfrac{1}{4}\sin 3\theta.$$

But these are the equations of an epicycloid of two cusps, *i.e.* the one in which $a = 2\,b$, $b = \frac{1}{4}$; cf. *Introduction to the Calculus*, p. 274.

EXERCISE

If the band is a complete circle and a point-source of light is situated on the circumference, draw accurately a figure showing the reflected rays and prove that their envelope is a cardioid.

4. Critique of the Method. The method set forth in § 1 has been given without any restriction on the functions, and it is easy to see that exceptions occur. Thus if equation (1) be solved for α:

$$\alpha = x \pm \sqrt{1 - y^2}.$$

and $f(x, y, \alpha)$ be written in the form

$$f(x, y, \alpha) = \alpha - x \mp \sqrt{1 - y^2},$$

the equation $\partial f/\partial\alpha = 0$ now reduces to $1 = 0$, and there is trouble. And yet

$$\alpha - x \mp \sqrt{1 - y^2} = 0$$

is just as much the equation of the family of circles as is equation (1), § 1.

A sufficient condition for the applicability of the method, and a condition which covers the most important cases which arise in practice, is given by the following theorem.

THEOREM. *Let $f(x, y, \alpha)$ be continuous, together with its derivatives of the first two orders, in the neighborhood of the point (x_0, y_0, α_0). Let*

(1) $$f(x_0, y_0, \alpha_0) = 0, \qquad f_\alpha(x_0, y_0, \alpha_0) = 0,$$

and let

(2) $$\begin{vmatrix} f_x & f_y \\ f_{x\alpha} & f_{y\alpha} \end{vmatrix} \neq 0, \qquad f_{\alpha\alpha} \neq 0,$$

in this point. Then the equations

(3) $$f(x, y, \alpha) = 0, \qquad f_\alpha(x, y, \alpha) = 0$$

define a curve which is tangent to each curve of the family

(4) $$f(x, y, \alpha) = 0$$

in the neighborhood of the point in question.

From the theorem of Chapter V, § 12, relating to implicit functions, it follows, since the determinant (2) is the Jacobian of the functions f and f_α, that equations (3) admit a simultaneous solution of the form

(5) $$x = \phi(\alpha), \qquad y = \psi(\alpha),$$

where $\phi(\alpha)$ and $\psi(\alpha)$ are continuous, together with their first derivatives, in the neighborhood of the point $\alpha = \alpha_0$, and where

$$\phi(\alpha_0) = x_0, \qquad \psi(\alpha_0) = y_0.$$

Moreover, the derivatives of ϕ and ψ do not both vanish. For, equations (3) become identities when x and y are replaced by $\phi(\alpha)$ and

$\psi(\alpha)$ respectively. Hence, on differentiating with respect to α, we have:

(6)
$$\begin{cases} f_x\phi' + f_y\psi' + f_\alpha = 0 \\ f_{x\alpha}\phi' + f_{y\alpha}\psi' + f_{\alpha\alpha} = 0 \end{cases}$$

The second equation proves that ϕ' and ψ' cannot both vanish, since by (2) $f_{\alpha\alpha} \neq 0$. Moreover, since $f_\alpha = 0$, the first equation reduces to

(7)
$$f_x\phi' + f_y\psi' = 0.$$

Thus equations (5) define a curve having a continuously turning tangent, and its slope is seen from (7) to be

$$\frac{\psi'(\alpha)}{\phi'(\alpha)} = -\frac{f_x}{f_y}.$$

On the other hand, since f_x and f_y cannot both vanish because of (2), equation (4) defines, for an arbitrary value of α, a curve having a continuously turning tangent, and its slope is

$$\frac{dy}{dx} = -\frac{f_x}{f_y}.$$

Hence, for an arbitrary value of α in the neighborhood of the value α_0, the curve (5) meets the curve (4) and since it has the same slope, is tangent to it.

Remark. It will be observed that the results of this paragraph are not merely destructive, but are primarily constructive. It is here not a question of a rigorous proof of a theorem, the truth of which no one doubts. The bare fact as to whether the method illustrated by the examples of § 1 has any standing is the first question at issue, and that question is answered by the theorem concerning implicit functions.

EXERCISES

1. Show that the family of curves

$$y = g(x, \alpha)$$

have an envelope in the neighborhood of the point (x_0, y_0), where

$$y_0 = g(x_0, \alpha_0),$$

provided that, at the point (x_0, α_0),

$$g_\alpha(x, \alpha) = 0, \qquad g_{x\alpha}(x, \alpha) \neq 0, \qquad g_{\alpha\alpha}(x, \alpha) \neq 0.$$

2. Show that the family of planes

$$z - z_0 = \alpha (x - x_0) + \beta (y - y_0), \qquad \Omega (\alpha, \beta) = 0,$$

envelop a cone with its vertex at (x_0, y_0, z_0), the generators being determined with the aid of the further equation

$$\frac{x - x_0}{\Omega_\alpha} = \frac{y - y_0}{\Omega_\beta}.$$

State precisely the conditions of continuity (including the existence of derivatives) you impose on the function Ω, and show that a sufficient condition can be formulated as the relation:

$$\Omega_{\beta\beta}\Omega_\alpha^2 - 2\Omega_{\alpha\beta}\Omega_\alpha\Omega_\beta + \Omega_{\alpha\alpha}\Omega_\beta^2 \neq 0.$$

CHAPTER IX

ELLIPTIC INTEGRALS

1. Origin and Definition of the Elliptic Integrals. The determination of the time of oscillation of a simple pendulum, *Introduction to the Calculus*, p. 373, is given by the equation

$$(1) \qquad t=\sqrt{\frac{l}{2g}}\int\frac{d\theta}{\sqrt{\cos\theta-\cos\alpha}}=\frac{1}{2}\sqrt{\frac{l}{g}}\int\frac{d\theta}{\sqrt{\sin^2\frac{\alpha}{2}-\sin^2\frac{\theta}{2}}}.$$

If we introduce a new variable of integration:

$$(2) \qquad \sin\frac{\theta}{2}=\sin\frac{\alpha}{2}\sin\phi, \qquad 0\leqq\theta\leqq\alpha, \qquad 0\leqq\phi\leqq\frac{\pi}{2},$$

we have:

$$\frac{1}{2}\cos\frac{\theta}{2}\,d\theta=\sin\frac{\alpha}{2}\cos\phi\,d\phi,$$

$$\sqrt{\sin^2\frac{\alpha}{2}-\sin^2\frac{\theta}{2}}=\sin\frac{\alpha}{2}\cos\phi,$$

$$\cos\frac{\theta}{2}=\sqrt{1-k^2\sin^2\phi}, \qquad k=\sin\frac{\alpha}{2}.$$

Hence

$$\frac{d\theta}{\sqrt{\sin^2\frac{\alpha}{2}-\sin^2\frac{\theta}{2}}}=\frac{2\,d\phi}{\sqrt{1-k^2\sin^2\phi}},$$

and the final formula for t is:

$$(3) \qquad t=\sqrt{\frac{l}{g}}\int_0^{\phi}\frac{d\phi}{\sqrt{1-k^2\sin^2\phi}}, \qquad k=\sin\frac{\alpha}{2}.$$

Here, ϕ can range through the interval $-\pi/2\leqq\phi\leqq\pi/2$; t is measured from the instant when the bob is lowest.

We are thus led to the function

$$(4) \qquad F(k,\phi)=\int_0^{\phi}\frac{d\phi}{\sqrt{1-k^2\sin^2\phi}}, \qquad 0<k<1,$$

which is known as the *Elliptic Integral of the First Kind* in *Legendre's* form. When $\phi = \pi/2$, we have

$$(5)\qquad K = \int_0^{\frac{\pi}{2}} \frac{d\phi}{\sqrt{1 - k^2 \sin^2 \phi}},$$

which is known as the *Complete Elliptic Integral of the First Kind.*

Arc of an Ellipse. The determination of the length of an arc of the ellipse

$$x = a \sin \phi, \qquad y = b \cos \phi$$

was found to be (*Introduction to the Calculus*, p. 414):

$$s = a \int_0^{\phi} \sqrt{1 - e^2 \sin^2 \phi}\, d\phi,$$

where e denotes the eccentricity of the ellipse.

We are thus led to the function

$$(6)\qquad E(k, \phi) = \int_0^{\phi} \sqrt{1 - k^2 \sin^2 \phi}\, d\phi, \qquad 0 < k < 1,$$

known as the *Elliptic Integral of the Second Kind* in *Legendre's* form. When $\phi = \pi/2$, we have:

$$(7)\qquad E = \int_0^{\frac{\pi}{2}} \sqrt{1 - k^2 \sin^2 \phi}\, d\phi,$$

known as the *Complete Elliptic Integral of the Second Kind.**

These are all tabulated functions, and abbreviated tables are given in Peirce's *Tables.* It is, therefore, of practical value to learn how to refer to the above certain other integrals that arise in practice.

Jacobi's Form. A second form of the elliptic integrals is known as *Jacobi's* form and arises through the change of variable

$$x = \sin \phi, \qquad -\frac{\pi}{2} \leqq \phi \leqq \frac{\pi}{2}, \qquad -1 \leqq x \leqq 1.$$

Thus

* The Elliptic Integral of the Third Kind is

$$\Pi(k, n, \phi) = \int_0^{\phi} \frac{d\phi}{(1 + n \sin^2 \phi)\sqrt{1 - k^2 \sin^2 \phi}}.$$

(8) $F(k, \phi) = \int_0^x \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}}, \qquad K = \int_0^1 \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}};$

(9) $E(k, \phi) = \int_0^x \frac{\sqrt{1-k^2x^2}}{\sqrt{1-x^2}}\,dx, \qquad E = \int_0^1 \frac{\sqrt{1-k^2x^2}}{\sqrt{1-x^2}}\,dx.$

The Complementary Modulus. The constant, or parameter, k, is known as the *modulus* of the integral; and k', defined by the equation

(10) $$k^2 + k'^2 = 1, \qquad 0 < k' < 1,$$

is called the *complementary modulus.* The corresponding value of K is denoted by K'.

(11) $$K' = \int_0^{\frac{\pi}{2}} \frac{d\phi}{\sqrt{1 - k'^2 \sin^2 \phi}}.$$

The Most General Elliptic Integral. Any integral of the type

$$\int R(x, \sqrt{(1-x^2)(1-k^2x^2)})\,dx,$$

where $R(x, y)$ is a rational function of x and y, and the integrand, on being simplified, actually involves the radical, is called an *elliptic integral.* Moreover, the radicand may be any polynomial of degree three or four, with distinct roots.

2. Integrals Reducible to $F(k, \phi)$. Any integral of the form

$$\int \frac{dx}{\sqrt{G_4(x)}} \quad \text{or} \quad \int \frac{dx}{\sqrt{G_3(x)}},$$

where G is a polynomial of the degree indicated, having all its roots distinct, can be reduced to $F(k, \phi)$ and thus evaluated by the *Tables* in any numerical case.

I.—THE INTEGRAL $\int \frac{dx}{\sqrt{\pm(1-x^2)(1-k^2x^2)}}$

The standard form is

(1) $$\int_0^x \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}} = F(k, \phi),$$

$$x = \sin \phi, \qquad -1 \leqq x \leqq 1, \qquad -\frac{\pi}{2} \leqq \phi \leqq \frac{\pi}{2}.$$

There are two other forms corresponding to the case that the roots of $G_4(x)$ are ± 1, $\pm 1/k$, namely:

$$(2) \qquad \int_{1/k}^{x} \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}}, \qquad \frac{1}{k} \leqq x < \infty;$$

$$(3) \qquad \int_{1}^{x} \frac{dx}{\sqrt{(x^2-1)(1-k^2x^2)}}, \qquad 1 \leqq x \leqq \frac{1}{k}.$$

In Cases (1) and (2) the radicand is positive in the intervals marked with a heavy line in Fig. 46. In Case (3) the intervals in

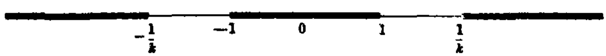


FIG. 46

which the radicand is positive are the supplementary intervals and are indicated by heavy lines in Fig. 47.

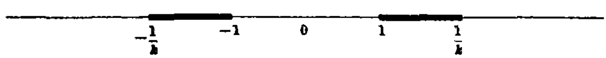


FIG. 47

It is obviously no restriction to consider only positive values for x, since, if x is negative, the substitution $x' = -x$ brings us back to the former case.

The integral (2) is reduced to the form (1) by the transformation:

$$t = \frac{1}{kx}, \qquad x = \frac{1}{kt}, \qquad 0 \leqq t \leqq 1;$$

$$(2') \qquad \int_{1/k}^{x} \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}} = K - \int_{0}^{t} \frac{dt}{\sqrt{(1-t^2)(1-k^2t^2)}}.$$

The transformation required for (3) is:

$$t = \frac{\sqrt{x^2-1}}{k'x}, \qquad x = \frac{1}{\sqrt{1-k'^2t^2}}; \qquad 0 \leqq t \leqq 1;$$

$$(3') \qquad \int_{1}^{x} \frac{dx}{\sqrt{(x^2-1)(1-k^2x^2)}} = \int_{0}^{t} \frac{dt}{\sqrt{(1-t^2)(1-k'^2t^2)}}.$$

The student should perform each time the actual analytic work of making the substitution.

Example. Required to compute the value of the integral

$$\int_1^{1.25} \frac{dx}{\sqrt{(x^2-1)(1-\frac{1}{4}x^2)}}.$$

Here, $k=\frac{1}{2}$, $k'=\frac{1}{2}\sqrt{3}$, and the upper limit of integration in the second integral (3′) is obtained by setting $x=\frac{5}{4}$ in the formula of transformation. Thus

$$t=\left.\frac{\sqrt{x^2-1}}{k'x}\right|_{x=\frac{5}{4}}=\frac{2\sqrt{3}}{5}=.6928,$$

and we need to know $F(\frac{1}{2}\sqrt{3},\ \phi)$, when $\sin\phi=.6928$.

Turning to Peirce's *Tables*, p. 122, and observing that α here has the value 60°, and $\phi=43°\ 51'$, we have to interpolate in the column headed 60°. The entries in that column are as follows:

$\phi=40°$,	.7436	.1076
$\phi=45°$,	.8512	

Since $3°\ 51'=3.850°$, interpolation by the rule of three, or first differences, gives the result .8265. But when the differences are so great, the last figure is meaningless, and even the third figure may be inexact by a unit (or possibly two).

Reference to the larger tables of Legendre * shows that the value to four places of decimals is .8260.

A More General Case. We are now in a position to evaluate the integrals

$$(i)\quad \int_0^x \frac{dx}{\sqrt{(A^2-x^2)(B^2-x^2)}},\qquad 0<A<B,\qquad 0\leqq x\leqq A;$$

$$(ii)\quad \int_B^x \frac{dx}{\sqrt{(A^2-x^2)(B^2-x^2)}},\qquad 0<A<B,\qquad B\leqq x\leqq \infty;$$

$$(iii)\quad \int_A^x \frac{dx}{\sqrt{(x^2-A^2)(B^2-x^2)}},\qquad 0<A<B,\qquad A\leqq x\leqq B.$$

The transformation

$$t=x/A,\qquad x=At,$$

carries these integrals respectively into the integrals (1), (2), (3), divided each time by B. Moreover, $k=A/B$.

* Legendre, *Traité des fonctions elliptiques*, Paris, 1826, vol. II, p. 292.

EXERCISES

Compute the values of the following integrals:

1. $\displaystyle\int_0^{-.5}\frac{dx}{\sqrt{(1-x^2)(1-\frac{1}{2}x^2)}}.$ *Ans.* -0.5356

2. $\displaystyle\int_2^{\infty}\frac{dx}{\sqrt{(1-x^2)(1-\frac{1}{4}x^2)}}.$

3. $\displaystyle\int_1^{8}\frac{dx}{\sqrt{(x^2-1)(1-0.01\,x^2)}}.$

Express the following integrals in terms of $F(k,\phi)$ or K.

4. $\displaystyle\int_0^{3}\frac{dx}{\sqrt{(25-x^2)(49-x^2)}}.$

5. $\displaystyle\int_2^{\infty}\frac{dx}{\sqrt{(3-2x^2)(5-3x^2)}}.$

6. $\displaystyle\int_3^{5}\frac{dx}{\sqrt{(x^2-9)(25-x^2)}}.$

7. $\displaystyle\int_3^{4}\frac{dx}{\sqrt{(x^2-9)(25-x^2)}}.$

II. — The Integral $\displaystyle\int\frac{dx}{\sqrt{\pm(1-x^2)(k'^2+k^2x^2)}}$

There are two cases here, as indicated in the figures. We give the transformation and the result, leaving the computation to the student.

$$(4)\qquad \int_0^{x}\frac{dx}{\sqrt{(1-x^2)(k'^2+k^2x^2)}}=K-\int_0^{t}\frac{dt}{\sqrt{(1-t^2)(1-k^2t^2)}},\qquad 0\leqq t\leqq 1,$$

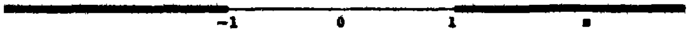
$$t=\sqrt{1-x^2},\qquad x=\sqrt{1-t^2}.$$

-1 0 x 1

Fig. 48a

$$(5)\qquad \int_1^{x}\frac{dx}{\sqrt{(x^2-1)(k'^2+k^2x^2)}}=\int_0^{t}\frac{dt}{\sqrt{(1-t^2)(1-k'^2t^2)}},\qquad 0\leqq t\leqq 1,$$

$$t=\frac{\sqrt{x^2-1}}{x},\qquad x=\frac{1}{\sqrt{1-t^2}}.$$



-1 0 1 x

Fig. 48b

To these integrals the following can be reduced by the transformation

$$y=\frac{x}{C}, \qquad x=Cy:$$

$$\text{(iv)}\quad \int_0^x \frac{dx}{\sqrt{(C^2-x^2)(A^2+B^2x^2)}}=\frac{1}{\sqrt{A^2+B^2C^2}}\int_0^y \frac{dy}{\sqrt{(1-y^2)(k'^2+k^2y^2)}},$$

$$0 \leqq y \leqq 1,$$

$$\text{(v)}\quad \int_C^x \frac{dx}{\sqrt{(x^2-C^2)(A^2+B^2x^2)}}=\frac{1}{\sqrt{A^2+B^2C^2}}\int_1^y \frac{dy}{\sqrt{(y^2-1)(k'^2+k^2y^2)}},$$

$$1 \leqq y \leqq \infty.$$

In each case,

$$k=\frac{BC}{\sqrt{A^2+B^2C^2}}, \qquad k'=\frac{A}{\sqrt{A^2+B^2C^2}}, \qquad A, B, C, \text{ positive}$$

III. — The Integral $\int \frac{dx}{\sqrt{(1+x^2)(1+k^2x^2)}}$

The upper limit of integration may be any positive number. We have:

$$\text{(6)}\quad \int_0^x \frac{dx}{\sqrt{(1+x^2)(1+k^2x^2)}}=\int_0^t \frac{dt}{\sqrt{(1-t^2)(1-k'^2t^2)}}, \qquad 0 \leqq t \leqq 1.$$

$$t=\frac{x}{\sqrt{1+x^2}}, \qquad x=\frac{t}{\sqrt{1-t^2}}.$$

Finally,

$$\text{(vi)}\quad \int_0^x \frac{dx}{\sqrt{(A^2+B^2x^2)(C^2+D^2x^2)}}=\frac{1}{AD}\int_0^y \frac{dy}{\sqrt{(1+y^2)(1+k^2y^2)}},$$

A, B, C, D, positive, $\qquad BC < AD, \qquad Dx=Cy, \qquad k=\frac{BC}{AD}.$

3. Continuation: $\int \frac{dx}{\sqrt{G_3(x)}}$.

Any polynomial of odd degree with real coefficients has at least one real root. Let $x=c$ be such a root of $G_3(x)$. Then the integral

$$\int_{x_0}^{x_1} \frac{dx}{\sqrt{G_3(x)}}$$

can be referred to an integral already in § 2 or § 4 by means of the transformation

$$x - c = z^2 \qquad \text{or} \qquad = -z^2, \qquad\qquad 0 \leqq z,$$

according as the values of x between x_0 and x_1 make $x - c$ positive or negative.

Example. Consider the integral:

$$\int_0^1 \frac{dx}{\sqrt{-(x-1)(x-2)(x-3)}}.$$

Let $x - 1 = -z^2$. The integral thus goes over into

$$\int_1^0 \frac{-2\,dz}{\sqrt{(1+z^2)(2+z^2)}} = 2\int_0^1 \frac{dz}{\sqrt{(1+z^2)(2+z^2)}}.$$

The substitutions $x - 2 = -z^2$ and $x - 3 = -z^2$ would have led to other forms equally tractable.

4. The General Case, $\displaystyle\int \frac{dx}{\sqrt{G_4(x)}}$.

Let $G_4(x)$ be a polynomial of the 4th degree, whose roots or factors are all distinct. If $G_4(x)$ has a real root, $x = \alpha$, the transformation

(1) $$y = \frac{1}{x - \alpha}, \qquad\qquad x = \alpha + \frac{1}{y},$$

will carry the integral into an integral of the form treated in § 3, namely:

$$\int \frac{dy}{\sqrt{G_3(y)}}.$$

It remains, therefore, merely to discuss the case that

(2) $$G_4(x) = (x^2 + p_1 x + q_1)(x^2 + p_2 x + q_2),$$

$$0 < 4q_1 - p_1^2, \qquad 0 < 4q_2 - p_2^2,$$

the second factor not being identical with the first. Let

(3) $$x = y + h.$$

Then $\quad x^2 + p_1 x + q_1 = y^2 + p_1' y + q_1', \qquad x^2 + p_2 x + q_2 = y^2 + p_2' y + q_2',$

where $\quad q_1' = h^2 + p_1 h + q_1, \qquad\qquad q_2' = h^2 + p_2 h + q_2.$

Let us seek to determine h so that q_1' and q_2' will be equal:

$$q_1' = q_2', \qquad h^2 + p_1 h + q_1 = h^2 + p_2 h + q_2,$$

(4) $$(p_1 - p_2)h + q_1 - q_2 = 0, \qquad h = -\frac{q_1 - q_2}{p_1 - p_2}.$$

We see, then, that this is possible except in the case $p_1 = p_2$. But here we attain our ultimate end immediately by the substitution

$$y = x + \tfrac{1}{2} p_1 = x + \tfrac{1}{2} p_2.$$

Secondly, we can reduce the constant terms, q_1' and $q_2' (= q_1')$, to unity by setting

(5) $$y = \kappa z$$

and choosing * $$\kappa = \sqrt{q_1'} = \sqrt{q_2'}.$$

Thus

$$G_4(x) = \kappa^4 (z^2 + P_1 z + 1)(z^2 + P_2 z + 1), \qquad P_1 \neq P_2.$$

We now make the final transformation:

(6) $$t = \frac{z-1}{z+1}, \qquad z = \frac{1+t}{1-t}.$$

Thus, $$z^2 + P_1 z + 1 = \frac{(2 + P_1) + (2 - P_1)t^2}{(1-t)^2},$$

with a similar formula for the second factor. We observe that neither coefficient, $2 + P_1$ or $2 - P_1$, can vanish,† and moreover, that these coefficients are either both positive or both negative. For, the quadratic polynomial in z has no real roots.

Example. $$u = \int_3^5 \frac{dx}{\sqrt{(x^2 - 4x + 7)(x^2 - 6x + 13)}}.$$

Here, $h = 3$, and $x = y + 3$:

$$u = \int_0^2 \frac{dy}{\sqrt{(y^2 + 2y + 4)(y^2 + 4)}}.$$

Furthermore, $$\kappa = 2, \quad y = 2z,$$

* That $q_1' > 0$ is clear from the fact that otherwise the quadratic polynomial in y would admit a real root, and hence the polynomial in x would, too.

† If $2 - P_1$ were $= 0$, make the transformation $t' = 1/t$, and the contradiction used in the proof below follows.

$$u = \frac{1}{2}\int_0^1 \frac{dz}{\sqrt{(z^2+z+1)(z^2+1)}}.$$

Finally, $\quad z = \dfrac{1+t}{1-t}, \qquad dz = \dfrac{2\,dt}{(1-t)^2},$

$$u = \frac{1}{\sqrt{2}}\int_{-1}^0 \frac{dt}{\sqrt{(3+t^2)(1+t^2)}}.$$

To compute this integral in terms of $F(k, \phi)$, set (§ 2, (6))

$$\tau = \frac{-t}{\sqrt{1+t^2}}, \qquad t = \frac{-\tau}{\sqrt{1-\tau^2}}, \qquad k = \frac{1}{\sqrt{3}}, \qquad k' = \sqrt{\frac{2}{3}}:$$

$$u = \frac{1}{\sqrt{6}}\int_0^{1/\sqrt{2}} \frac{d\tau}{\sqrt{(1-\tau^2)(1-\frac{2}{3}\tau^2)}}.$$

5. Computation by Series. We have already seen how the functions $F(k, \phi)$ and $E(k, \phi)$ can be computed by infinite series; *Introduction to the Calculus*, pp. 414, 416. These series do not, however, converge rapidly when k is nearly unity. In this case, a transformation can be made (*Landen's Transformation*, § 6) whereby either (*a*) k will be replaced by a smaller value, k_1, and thus the new series will become available for practical use; or (*b*) k can be replaced by a still larger value, so near to unity that it may be set $= 1$ in the integral, and then the latter can be computed by means of the indefinite integral.

6. Landen's Transformation. We give the transformation without *motif.** Starting with the integral

$$(1) \qquad F(k, \phi) = \int_0^\phi \frac{d\phi}{\sqrt{1-k^2\sin^2\phi}}$$

we introduce a new variable of integration, ψ, by the equation

$$(2) \qquad \sin(2\psi - \phi) = k\sin\phi,$$

* The mathematicians of the eighteenth century were men of great resourcefulness in formal work, and many of the leading results in the theory of the elliptic integrals and functions were deduced by inspiration rather than by reasoning. On the other hand, the modern theory of transformation of the elliptic transcendents is too complex to admit of a brief description.

or its equivalent,

$$(3)\qquad \tan\phi = \frac{\sin 2\psi}{k+\cos 2\psi}, \qquad \text{or} \qquad \tan(\psi-\phi) = \frac{1-k}{1+k}\tan\psi.$$

From the first of the equations (3) we have:

$$(4)\qquad d\phi = 2\,\frac{1+k\cos 2\psi}{1+2k\cos 2\psi+k^2}\,d\psi.$$

From (4) it appears that $d\psi/d\phi$ is always positive, and so, as ϕ increases from 0 to π, the determination of ψ with which we are concerned will increase from 0 to $\pi/2$.

Furthermore, from the first of equations (3), $\sin^2\phi$ can be computed, and thus we find:

$$(5)\qquad \sqrt{1-k^2\sin^2\phi} = \frac{1+k\cos 2\psi}{\sqrt{1+2k\cos 2\psi+k^2}}.$$

From (4) and (5) we have:

$$(6)\qquad \frac{d\phi}{\sqrt{1-k^2\sin^2\phi}} = \frac{2\,d\psi}{\sqrt{1+2k\cos 2\psi+k^2}} = \frac{2}{1+k}\,\frac{d\psi}{\sqrt{1-k_1^2\sin^2\psi}},$$

$$(7)\qquad k_1^2 = \frac{4k}{(1+k)^2}, \qquad k = \frac{1-\sqrt{1-k_1^2}}{1+\sqrt{1-k_1^2}}.$$

On integrating equation (6) we find:

$$\int_0^\phi \frac{d\phi}{\sqrt{1-k^2\sin^2\phi}} = \frac{2}{1+k}\int_0^\psi \frac{d\psi}{\sqrt{1-k_1^2\sin^2\psi}},$$

where the limits of integration, ϕ and ψ, are connected by (2), or either of the forms (3).

To sum up, then, we have:

$$(8)\qquad \begin{cases} F(k,\phi) = \dfrac{2}{1+k}\,F(k_1,\psi) \\[2ex] k_1 = \dfrac{2\sqrt{k}}{1+k}, \qquad \sin(2\psi-\phi) = k\sin\phi. \end{cases}$$

The new modulus, k_1, is greater than k, but less than 1. For, first, if

$$\frac{2\sqrt{k}}{1+k} > k, \qquad \text{then} \qquad \frac{4k}{(1+k)^2} > k^2 \qquad \text{and} \qquad 4 > k(1+k)^2.$$

This last inequality is true, since $0<k<1$; and now, starting with it, we can retrace our steps. In a similar manner it is shown that $k_1<1$.

If k is nearly $=1$, a few repetitions of the transformation (7) will lead to a function $F(k_n, \phi_n)$, whose modulus k_n is so nearly unity that it may be replaced by 1 and the integral thus evaluated. Since

$$\frac{2}{1+k}=\frac{k_1}{\sqrt{k}},$$

we have:

$$F(k, \phi)=k_n\sqrt{\frac{k_1 k_2 \cdots k_{n-1}}{k}}\,F(k_n, \phi_n).$$

On setting $k_n=1$, we find:

$$F(k_n, \phi_n)=\int_0^{\phi_n}\frac{d\phi}{\cos\phi}=\log\tan\left(\frac{\pi}{4}+\frac{\phi_n}{2}\right).$$

For a detailed study of a numerical case, cf. Byerly, *Integral Calculus*, 2d ed., chapter on Elliptic Integrals.

Reducing the Modulus. The transformation (7) can be applied in the opposite sense, and thus the given integral is referred to one with smaller modulus. The formulas now become:

$$(9)\qquad \begin{cases} F(k_1, \psi)=\dfrac{1+k}{2}F(k, \phi), \\ k=\dfrac{1-\sqrt{1+k_1^2}}{1+\sqrt{1+k_1^2}}, \qquad \tan(\phi-\psi)=\dfrac{1-k}{1+k}\tan\psi. \end{cases}$$

Here, k_1 and ψ are given, and k and ϕ are computed from the second line of (9). The student may find it convenient to rewrite (9), interchanging ϕ with ψ and k with k_1. A numerical example is worked in detail in Byerly's book, *l.c.*

After one or two applications of the transformation (9), it may be well to finish the computation by using the series.

Integrals of the Second Kind, $\int\sqrt{1-k^2\sin^2\phi}\,d\phi$. Landen's transformation can be applied to these, too, and thus the computation carried through; cf. Byerly, *l.c.* An excellent treatment of this subject, including also the rectification of the hyperbola and the lemniscate, and the complanation of the central quadrics, is found in Schlömilch, *Compendium der höheren Analysis*, vol. 2, 2d ed.

7. The Elliptic Functions. If we set

$$u = \int_0^x \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}}, \qquad -1 \leqq x \leqq 1,$$

the equation represents u as a function of x. The inverse function, x, regarded as a function of u, is called the *sine amplitude of* u and is written

$$x = \sin \operatorname{am} u \qquad \text{or} \qquad x = \operatorname{sn} u.$$

Two other functions are defined by the equations:

$$\sqrt{1-x^2} = \cos \operatorname{am} u \qquad \text{or} \qquad \operatorname{cn} u;$$

$$\sqrt{1-k^2x^2} = \Delta \operatorname{am} u \qquad \text{or} \qquad \operatorname{dn} u$$

(read: "delta amplitude of u" or "d n u.") These functions are known as *Elliptic Functions*, and any rational function of them is also called an elliptic function. For a brief treatment of them, cf. Byerly, *l.c.* A more extended study is found in Pierpont's *Functions of a Complex Variable*, Chapters X–XII. Cf. also, both for the elliptic integrals and the elliptic functions, Schlömilch, *l.c.*

CHAPTER X

INDETERMINATE FORMS *

1. The Limit $\frac{0}{0}$. Let two functions, $f(x)$ and $F(x)$, be continuous in the interval

(1) $$a \leqq x \leqq b,$$

and let

(2) $$f(a) = 0, \qquad F(a) = 0$$

The ratio of these functions,

(3) $$\frac{f(x)}{F(x)},$$

will not be defined when $x = a$, since it takes on the form $0/0$, and division by 0 is impossible. Nevertheless, the ratio (3) may approach a limit when x approaches a, and this is, in fact, usually the case in practice. For example,

(*i*) $$\frac{x-a}{x^2-a^2} = \frac{1}{x+a}, \qquad \lim_{x \doteq a} \frac{x-a}{x^2-a^2} = \lim_{x \doteq a} \frac{1}{x+a} = \frac{1}{2a};$$

(*ii*) $$\lim_{x \doteq 0} \frac{\sin x}{x} = 1.$$

Suppose, furthermore, that each function has a continuous derivative at every interior point of the interval, and that $F'(x) \neq 0$ there:

(4) $$F'(x) \neq 0, \qquad a < x < b.$$

By the Law of the Mean, Chap. V, § 2,

$$f(a+h) - f(a) = hf'(X), \qquad F(a+h) - F(a) = hF'(X'),$$
$$0 < h < b - a,$$

where $\quad a < X < a + h, \qquad a < X' < a + h.$

* This subject was formerly made much of in a first course in the Calculus, doubtless because it yielded a vast fund of problems in differentiation. But we have not yet needed it, nor shall we find an application for the results till we take up Improper Integrals in Chapter XIX.

Since (2) is true, we have:

$$(5)\qquad \frac{f(a+h)}{F(a+h)}=\frac{f'(X)}{F'(X')}.$$

If, now, these derivatives exist and are continuous at the point $x=a$, too; and if $F'(a)\neq 0$, then*

$$(6)\qquad \lim_{x\doteq a}\frac{f(x)}{F(x)}=\lim_{h\doteq 0}\frac{f'(X)}{F'(X')}=\frac{f'(a)}{F'(a)}.$$

For example,

$$(7)\qquad \lim_{x\doteq 0}\frac{e^x-1}{x}=\frac{e^0}{1}=1.$$

But if, as in the case of

$$(8)\qquad \frac{1-\cos x}{x^2},\qquad a=0,$$

$f'(a)$ and $F'(a)$ both vanish, equation (6) breaks down, nor can we do anything with (5) since we do not know how X and X' vary relatively to each other. This case can be dealt with as follows.

Generalized Law of the Mean. *If $f(x)$ and $F(x)$ are continuous throughout the interval $a\leqq x\leqq b$ and each has a derivative at all interior points of the interval, and if, moreover, the derivative $F'(x)$ does not vanish within the interval; then, for some value $x=X$ within this interval,*

$$(9)\qquad \frac{f(b)-f(a)}{F(b)-F(a)}=\frac{f'(X)}{F'(X)},\qquad a<X<b.$$

* In this case the result can be obtained at once, since

$$\frac{f(x)}{F(x)}=\frac{f(a+h)-f(a)}{h}\Big/\frac{F(a+h)-F(a)}{h},$$

and the limit of the right-hand side is seen to be $f'(a)/F'(a)$. This is known as "l'Hospital's Rule," dating from 1696.

The limit is also called the "true value" of the "indeterminate form" $f(x)/F(x)$ for $x=a$. Both terms are based on a false conception. In the early days of the Calculus mathematicians thought of the fraction as really having a value when $x=a$, only the value cannot be computed because the form of the fraction eludes us. This is wrong. Division by 0 is not a process which we define in Algebra. It is convenient, however, to retain the term *indeterminate form* as applying to such expressions as the above and others considered in this chapter, which for a certain value of the independent variable cease to have a meaning, but which approach a limit when the independent variable converges toward the exceptional value.

Proof. *The function* *

$$\phi(x)=\frac{f(b)-f(a)}{F(b)-F(a)}[F(x)-F(a)]-[f(x)-f(a)]$$

satisfies all the conditions of Rolle's Theorem, *Introduction*, p. 430, and hence its derivative,

$$\phi'(x)=\frac{f(b)-f(a)}{F(b)-F(a)}F'(x)-f'(x),$$

must vanish for a value of x within the interval. Hence

$$\frac{f(b)-f(a)}{F(b)-F(a)}F'(X)-f'(X)=0, \qquad a<X<b.$$

By hypothesis, $F'(x)$ is never 0 in the interval. Consequently we are justified in dividing through by $F'(X)$, and thus (9) is established.

The Limit $\frac{0}{0}$, *Concluded.* We can now deduce a more general rule for determining the limit of the function (3). Applying (9) to an arbitrary sub-interval (a, x), when $a<x<b$, and remembering that $f(a)=0$ and $F(a)=0$, we see that

(10) $$\frac{f(x)}{F(x)}=\frac{f'(X)}{F'(X)}, \qquad a<X<x,$$

where now we have the *same* X in numerator and denominator.

When x approaches a, X will also approach a. Hence, if $f'(x)/F'(x)$ approaches a limit, $f'(X)/F'(X)$ will approach the same limit, and so will its equal, $f(x)/F(x)$. Thus we have:

(I) $$\lim_{x\doteq a}\frac{f(x)}{F(x)}=\lim_{x\doteq a}\frac{f'(x)}{F'(x)}.$$

If, then, it turns out on differentiating that $f'(a)=0$ and $F'(a)=0$, we can differentiate again, and so on.

Example.

$$\lim_{x\doteq 0}\frac{x-\sin x}{x^3}=\lim_{x\doteq 0}\frac{1-\cos x}{3x^2}=\lim_{x\doteq 0}\frac{\sin x}{6x}=\frac{1}{6}.$$

* We can divide by $F(b)-F(a)$, since

$$F(b)-F(a)=(b-a)F'(Y), \qquad a<Y<b,$$

and neither factor on the right is 0.

Remark. The Theorem can be extended to include the case that $f'(x)/F'(x)$ becomes positively infinite or negatively infinite. The function $f(x)/F(x)$ then becomes positively infinite or negatively infinite.

Moreover, instead of approaching a from above, x may approach a from below, or x may become positively infinite or negatively infinite.

EXERCISES

Determine the following limits:

1. $\lim\limits_{x\to -1} \dfrac{\sin \pi x}{1+x}$.
2. $\lim\limits_{x\to \frac{\pi}{4}} \dfrac{\tan x - 1}{\frac{\pi}{4} - x}$.
3. $\lim\limits_{x\to 0} \dfrac{a^x - b^x}{x}$.
4. $\lim\limits_{x\to 0} \dfrac{1-\cos x}{\tan^2 x}$.
5. $\lim\limits_{x\to 0} \dfrac{\sin x - x}{\tan x - x}$.
6. $\lim\limits_{x\to +\infty} \dfrac{\cot^{-1} x}{\csc^{-1} x}$.

2. The Limit $\frac{\infty}{\infty}$. Consider the fraction

$$\frac{f(x)}{F(x)}, \tag{1}$$

where $f(a) = \infty$ and $F(a) = \infty$. We wish to determine

$$\lim_{x\to a} \frac{f(x)}{F(x)}. \tag{2}$$

Simple cases, like

$$\lim_{n=\infty} \frac{2n-1}{3n+1}, \qquad \lim_{x=\infty} \frac{x}{\sqrt{x^2-1}},$$

are dealt with directly by obvious algebraic reductions; *cf. Introduction*, p. 29. In less transparent cases the following theorem often makes possible the evaluation.

THEOREM. *Let $f(x)$ and $F(x)$ be defined in the interval $a < x \leqq b$, and let*

$$f(a) = \infty, \qquad F(a) = \infty.$$

Let $f'(x)$ and $F'(x)$ exist at each point of the above interval and let $F'(x)$ be different from 0 there. If $f'(x)/F'(x)$ approaches a limit as x approaches a, then $f(x)/F(x)$ also approaches a limit, and these limits are equal:

$$\lim_{x\to a} \frac{f(x)}{F(x)} = \lim_{x\to a} \frac{f'(x)}{F'(x)}.$$

For convenience, let $a=+\infty$, the interval then becoming $g \leqq x < \infty$. By the Generalized Law of the Mean, § 1, we have

$$\frac{f(x)-f(x')}{F(x)-F(x')}=\frac{f'(X)}{F'(X)}, \qquad g < x' < x, \qquad x' < X < x.$$

Hence

$$(3) \qquad \frac{f(x)}{F(x)}=\lambda\frac{f'(X)}{F'(X)}, \qquad \lambda=\frac{1-F(x')/F(x)}{1-f(x')/f(x)}.$$

Now let x and x' both become infinite; but let x increase so much more rapidly that

$$\lim_{x=\infty,\, x'=\infty}\frac{f(x')}{f(x)}=0, \qquad \lim_{x=\infty,\, x'=\infty}\frac{F(x')}{F(x)}=0.$$

Then $\lim\lambda=1$. Now, X becomes infinite, and hence the whole right-hand side of the first equation (3) approaches the limit which $f'(x)/F'(x)$ by hypothesis approaches. Thus the theorem is proved.*

If $f'(x)/F'(x)$ becomes positively infinite, or negatively infinite, the same is true of $f(x)/F(x)$.

If a is a finite point, the same reasoning still holds, with obvious modifications in details. Or, this case can be referred directly to the above by means of such a substitution as

$$y=1/(x-a), \qquad x=a+1/y.$$

This theorem has the same advantage as that of § 1, namely, that, if we do not get a result after the first pair of differentiations, we may differentiate again and again. If, after k repetitions, we do get a result, then the original ratio approaches this same limit.

Example 1. $\lim\limits_{x=\infty}\dfrac{x^n}{e^x}$.

If $n \leqq 0$, the limit is obviously 0, since the numerator remains finite and the denominator becomes infinite. If, however, $n>0$, we see that the above ratio approaches 0 provided $n \leqq 1$. If $n>1$, a finite number of repetitions will lead to a ratio whose limit is 0, and thus the given ratio approaches 0 for *any fixed* value of n.

Example 2. $\lim\limits_{x=\infty}\dfrac{(\log x)^a}{x^\beta}, \qquad 0<\alpha, \qquad 0<\beta.$

If $\alpha=1$, we have

$$\lim_{x=\infty}\frac{\log x}{x^\beta}=\lim_{x=\infty}\frac{1}{\beta x^\beta}=0.$$

* The theorem is due to Cauchy, who gave a proof under certain restrictions. The complete proof, given above, is due to the Austrian mathematician Stolz.

If $\alpha \neq 1$, write

$$\frac{(\log x)^\alpha}{x^\beta} = \left[\frac{\log x}{x^{\beta/\alpha}}\right]^\alpha.$$

Since the variable standing within the brackets approaches 0, and since α is positive, the whole right-hand side approaches 0, and thus the original variable approaches 0 in all cases.*

3. The Limit $0 \cdot \infty$. If $f(x)$ approaches 0 and $\phi(x)$ becomes infinite, the limit approached by the product may often be determined by writing

$$f(x)\,\phi(x) = \frac{f(x)}{1/\phi(x)} \qquad \text{or} \qquad = \frac{\phi(x)}{1/f(x)},$$

thus throwing the variable into the form discussed in § 1 or § 2.

Example. $\lim\limits_{x \doteq 0} x \log x$.

$$x \log x = \frac{\log x}{x^{-1}}, \qquad \lim_{x \doteq 0} (x \log x) = \lim_{x \doteq 0} \frac{\log x}{x^{-1}} = \lim_{x \doteq 0} (-x) = 0.$$

4. The Limits $0^0, 1^\infty, \infty^0, \infty - \infty$. The function $f(x)^{\phi(x)}$ can be written in the form

$$f(x)^{\phi(x)} = e^{\phi(x) \log f(x)}.$$

When one factor in this last exponent approaches 0 and the other becomes infinite, the limit of the exponent is of the type considered in § 3. Thus we are led to the limits which may be symbolized as 0^0, 1^∞, ∞^0.

Example. $\lim\limits_{x \doteq 0} x^x$.

Since $\lim\limits_{x \doteq 0} (x \log x) = 0$ by § 3,

$$\lim_{x \doteq 0} x^x = \lim_{x \doteq 0} e^{x \log x} = 1.$$

The limit of $f(x) - \phi(x)$, where $f(x)$ and $\phi(x)$ both become infinite with the same sign, is usually best treated by special methods.

Example. $\lim\limits_{x \doteq \infty} \{\sqrt{x^2+1} - x\}$.

Write $\sqrt{x^2+1} - x = \dfrac{\{\sqrt{x^2+1} - x\}\{\sqrt{x^2+1} + x\}}{\sqrt{x^2+1} + x} = \dfrac{1}{\sqrt{x^2+1} + x},$

* A thorough appreciation of the meaning of the graph of the function $y = x^x$, *Introduction to the Calculus*, p. 160, is important in the study of the present chapter.

and the limit is obvious. Or, the method of series may be used:

$$\sqrt{x^2+1} = x\left(1+\frac{1}{x^2}\right)^{\frac{1}{2}} = x\left\{1+\frac{1}{2}\frac{1}{x^2}+\cdots\right\}, \text{ etc.}$$

EXERCISES ON CHAPTER X

Determine in the most convenient manner possible each of the following limits.

1. $\lim\limits_{x=\infty} xe^{-x^2}$.
2. $\lim\limits_{x=\infty} x^n e^{-x^2}$.
3. $\lim\limits_{x=\infty} x^n e^{-ax}$, $0 < a$.
4. $\lim\limits_{x=\infty} \dfrac{\log x}{1+x+x^2}$.
5. $\lim\limits_{x=\infty} \dfrac{(\log x)^n}{\sqrt{1+x+x^2+x^5}}$.
6. $\lim\limits_{x=\infty} x^n e^{-x} \log x$. Suggestion: $x^n e^{-x} \log x = \left(\dfrac{x^{n+1}}{e^x}\right)\left(\dfrac{\log x}{x}\right)$.
7. $\lim\limits_{x=\infty} x^n e^{-x} (\log x)^2$.
8. $\lim\limits_{x=\infty} \dfrac{x^\alpha}{(\log x)^\beta}$, $\begin{cases} 0 < \alpha, \\ 0 < \beta. \end{cases}$
9. $\lim\limits_{x=\infty} e^{-x} \cot^{-1} x$.
10. $\lim\limits_{x=\infty} \dfrac{e^{-\frac{1}{x^2}}}{x^n}$.
11. $\lim\limits_{x=\frac{\pi}{2}} (\sin x)^{\tan x}$.

CHAPTER XI

LINE INTEGRALS AND GREEN'S THEOREM. FLOW OF HEAT

1. Work. We have defined the *work*, W, done on a particle by a variable force, F, when the particle moves along a straight line and the force acts along the same line, by means of the integral (*Introduction to the Calculus*, p. 338):

$$W = \int_a^b F\,dx. \tag{1}$$

Here, F may be any continuous function of x, positive or negative, and thus W is also a signed quantity. The interval throughout which the particle is displaced may be variable. Thus if ξ be any point of the interval $a \leqq x \leqq b$, and the particle be displaced from a to ξ, the work done will be

$$W = \int_a^\xi F\,dx.$$

Definite Integral as Function of Upper Limit of Integration. We have here before us a first example of a function represented by a definite integral, the upper limit of integration being the independent variable. Let $f(x)$ be continuous in the interval $a \leqq x \leqq b$, and let ξ be any point of this interval. Then

$$\int_a^\xi f(x)\,dx$$

is a *function of* ξ, which we will denote by $\phi(\xi)$. If we change the notation, denoting the variable of integration by t and the upper limit by x, we have: *

$$\phi(x) = \int_a^x f(t)\,dt.$$

* The notation $\int_a^x f(x)\,dx$ should be avoided till the student is thoroughly conscious of the different meanings of the letter x in this expression.

The function $\phi(x)$ thus represented or defined admits a derivative, obtained as follows. Since

$$\phi(x_0+h)-\phi(x_0)=\int_{x_0}^{x_0+h} f(t)\,dt,$$

we have, on applying the Law of the Mean,

$$\int_{x_0}^{x_0+h} f(t)\,dt = hf(x_0+\theta h), \qquad 0<\theta<1.$$

Hence

$$\frac{\phi(x_0+h)-\phi(x_0)}{h}=f(x_0+\theta h) \quad \text{and} \quad \lim_{h\doteq 0}\frac{\phi(x_0+h)-\phi(x_0)}{h}=f(x_0)$$

Thus we have proved the theorem that

$$\frac{d}{dx}\int_a^x f(t)\,dt = f(x).$$

EXERCISE

Prove that $\displaystyle \frac{d}{dx}\int_x^b f(t)\,dt = -f(x).$

2. Continuation: Curved Paths. Suppose the particle describes a curved path C in a plane, and that the force, $\mathfrak{F}$, varies in magnitude and direction in any continuous manner. What will be the work done in this case?

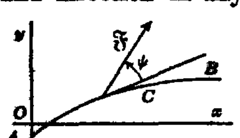


FIG. 49

Suppose the path C is a right line and the force, though oblique to the line, is constant in magnitude and direction; Fig. 50. Resolve the force into its two components along the line and normal to it. Surely, we must lay down our definition of work so that the work done by $\mathfrak{F}$ is equal to the sum of the works of the component forces. Now, the work done by the component along the line has already been defined, namely, $Fl\cos\psi$, where $F=|\mathfrak{F}|$ is the intensity of the force.

It is an essential part of the idea of *work* that the force overcomes resistance through distance (or is overcome through distance). Now, the normal component does neither; it merely sidles off and sidesteps the whole question. It is natural, therefore, to define it as doing no work. Thus we arrive at our final definition: The work done by $\mathfrak{F}$ in the particular case in hand shall be

(1) $$W = Fl\cos\psi.$$

A second form of the expression on the right is as follows. Let X and Y be the components of $\mathfrak{F}$ along the axes. Let τ be the angle that the path AB makes with the positive axis of x. Then the projection of $\mathfrak{F}$ on AB is equal to the sum of the projections of X and Y on AB, or

$$F\cos\psi = X\cos\tau + Y\sin\tau.$$

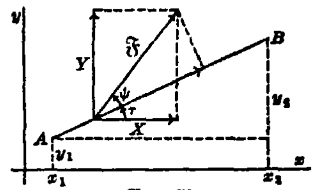


Fig. 50

On the other hand,

$$x_2 - x_1 = l\cos\tau, \qquad y_2 - y_1 = l\sin\tau.$$

Hence

(2) $$W = X(x_2 - x_1) + Y(y_2 - y_1).$$

The General Case. If C be any regular curve, divide it into n arcs by the points $s_0 = 0, s_1, \cdots, s_{n-1}, s_n = l$. Let $\mathfrak{F}'_k$ be the value of $\mathfrak{F}$ at an arbitrary point of the k-th arc, and let ψ'_k be the angle from the chord (s_{k-1}, s_k) to the vector $\mathfrak{F}'_k$. Then the sum

$$\sum_{k=1}^{n} F'_k \cos\psi'_k\, l_k,$$

where l_k denotes the length of the chord, gives us approximately what we should wish to understand by the work, in view of our physical feeling for this quantity. The limit of this sum, when the longest l_k approaches 0, shall be defined as the work, or

(3) $$W = \lim_{n=\infty} \sum_{k=1}^{n} F'_k \cos\psi'_k\, l_k.$$

Since $$\lim \frac{l_k}{\Delta s_k} = 1,$$

it is clear that the above limit is the same as

$$\lim_{n=\infty} \sum_{k=1}^{n} F_k \cos\psi_k\, \Delta s_k = \int_0^l F\cos\psi\, ds.$$

For, the conditions of Duhamel's Theorem are fulfilled if

$$\alpha_k = F_k\cos\psi_k\,\Delta s_k, \qquad \beta_k = F'_k\cos\psi'_k\, l_k,$$

since $$\frac{\beta_k}{\alpha_k} = \left(\frac{F'_k}{F_k}\right)\left(\frac{\cos\psi'_k}{\cos\psi_k}\right)\left(\frac{l_k}{\Delta s_k}\right), \qquad \text{and thus} \qquad \lim\frac{\beta_k}{\alpha_k} = 1.$$

We have, then, as the expression for the work,

$$(4)\qquad W=\int_{0}^{l} F\cos\psi\,ds.$$

A second formula for the work is obtained from (2), namely,

$$(5)\qquad W=\lim_{n=\infty}\sum_{k=1}^{n}(X'_k\,\Delta x_k+Y'_k\,\Delta y_k).$$

This limit can also be expressed as an integral. Since

$$\lim\frac{\Delta x_k}{\Delta s_k\cos\tau_k}=1\qquad\text{and}\qquad\lim\frac{\Delta y_k}{\Delta s_k\sin\tau_k}=1,$$

we see that, on setting

$$\alpha_k=(X_k\cos\tau_k+Y_k\sin\tau_k)\,\Delta s_k,\qquad \beta_k=X'_k\,\Delta x_k+Y'_k\,\Delta y_k,$$

the conditions of Duhamel's Theorem are fulfilled, and hence the above limit has the value:

$$(6)\qquad \int_{0}^{l}(X\cos\tau+Y\sin\tau)\,ds\qquad\text{or}\qquad\int_{0}^{l}\left(X\frac{dx}{ds}+Y\frac{dy}{ds}\right)ds.$$

Thus the limit (5) is seen to exist, and to have for its value the integral (6). The limit (5) is an example of a *line integral*, and is expressed by the following notation, § 3:

$$(7)\qquad W=\int_{(a,\,b)}^{(a',\,b')}X\,dx+Y\,dy\qquad\text{or}\qquad\int_{C}X\,dx+Y\,dy.$$

The extension to three dimensions is immediate. Formula (4) requires no modification whatever. Formula (7) is replaced by the following:

$$(8)\quad W=\int_{(a,\,b,\,c)}^{(a',\,b',\,c')}X\,dx+Y\,dy+Z\,dz\qquad\text{or}\qquad\int_{C}X\,dx+Y\,dy+Z\,dz.$$

Example 1. To find the work done by gravity on a particle of mass m which moves from an initial point (x_0, y_0, z_0) to a final point (x_1, y_1, z_1) along an arbitrary twisted curve, C.

Let the axis of z be vertical and positive downwards. Then

$$X=0,\qquad Y=0,\qquad Z=mg;$$

$$W=\int_{C}X\,dx+Y\,dy+Z\,dz=\int_{z_0}^{z_1}mg\,dz=mg(z_1-z_0).$$

Hence the work done is equal to the product of the force by the difference in level, and depends only on the initial and final points, but not on the path joining them.

Example 2. Consider a field of force, corresponding to a force function, u (Chap. V, § 17). Then the components of the force which acts on a unit particle at any point of the field will be:

$$X=\frac{\partial u}{\partial x}, \qquad Y=\frac{\partial u}{\partial y}, \qquad Z=\frac{\partial u}{\partial z}.$$

Let the particle describe a curve C in the field, running from the point A to the point B. The work done by the field on the particle will be

$$W=\int\limits_C \frac{\partial u}{\partial x}dx+\frac{\partial u}{\partial y}dy+\frac{\partial u}{\partial z}dz=u|_C=u_B-u_A.$$

Hence the work done is equal to the change in value of the force function, taken along the path.

EXERCISES

1. A well is pumped out by a force pump which delivers the water at the mouth of a pipe which is fixed. Show that the work done is equal to the weight of the water initially in the well, multiplied by the vertical distance of the centre of gravity below the mouth of the pipe.

2. A meteor, which may be regarded as a particle, is attracted by the sun (considered at rest) and by all the rest of the matter in the solar system. It moves from a point A to a point B. Show that the work done on it by the sun is

$$W=Km\left(\frac{1}{r_1}-\frac{1}{r_0}\right),$$

where r_0 and r_1 represent the distances of A and B, respectively, from the sun, and K is the gravitational constant.

3. A straight wire carries a current, thus generating an electromagnetic field of force. The force which acts on a unit north pole is inversely proportional to the distance from the wire, and in the direction at right angles to the wire and to the perpendicular dropped on the wire. Find the work done on a unit north pole, when the latter describes a circle, the axis of which lies along the wire.

3. Line Integrals. The limits (3) and (5), § 2, are typical illustrations of how line integrals come into mathematics. Let S be a region of the (x, y)-plane, and let

$$C: \qquad x = f(t), \qquad y = \phi(t), \qquad t_0 \leqq t \leqq t_1$$

be a regular curve lying in S. Let $F(x, y, t)$ depend (continuously) on the point (x, y) in S and the point t of C. Divide C into n arcs by the points $s_0 = 0, s_1, \cdots, s_{n-1}, s_n = l$, where s denotes the arc, and l, the length of C. Let (x'_k, y'_k) be an arbitrary point of the k-th arc, (s_{k-1}, s_k), and let t''_k be any second point of the same arc. Then

$$(1) \qquad \lim_{n=\infty} \sum_{k=1}^{n} F(x'_k, y'_k, t''_k)\, \Delta s_k = \int_C F\, ds$$

is defined as the *line integral* of the function F along the curve C.

That this limit exists is clear from Duhamel's Theorem, since

$$(2) \qquad \lim_{n=\infty} \sum_{k=1}^{n} F(x_k, y_k, t_k)\, \Delta s = \int_0^l F\, ds$$

is the ordinary integral of $F(x, y, t)$, a continuous function of s.

It is particularly to be remarked that, in the definition (1), Δs_k is not a signed quantity, but is essentially positive. Thus the value of the line integral (1) does not depend on the sense of integration along C. We might equally well integrate in the opposite sense; the result would be the same. On the other hand, the line integrals presently to be defined are signed quantities. Reversal of the sense of integration along C reverses the signs of these integrals.*

Definition of the Line Integral $\int_C P\,dx + Q\,dy$.

Let P be a function of (x, y), continuous throughout S. Let C be given as before, and let C be divided into n arcs by the points (x_k, y_k). Form the sum:

$$(3) \qquad \sum_{k=1}^{n} P(x'_k, y'_k)\, \Delta x_k, \qquad \Delta x_k = x_k - x_{k-1},$$

* It might seem that the integral for the work, $W = \int_C F \cos\psi\, ds$, is an exception, since reversing the sense in which the particle describes C reverses the sign of the work. But when the sense is reversed, ψ is replaced by its supplement $\pi - \psi$, and thus the sign of $\cos\psi$ is reversed.

where (x'_k, y'_k) is any point of the k-th arc. Then this sum approaches a limit when n becomes infinite, the longest arc approaching 0. For, let $x = \omega(s)$, where s_k increases with k. Then

$$\Delta x_k = \omega(s_k) - \omega(s_{k-1}) = \Delta s_k \cos \tau''_k.$$

Now

$$(4) \qquad \lim_{n=\infty} \sum_{k=1}^{n} P(x_k, y_k) \cos \tau_k \, \Delta s_k = \int_0^l P \cos \tau \, ds$$

is the ordinary integral of $P \cos \tau$, a continuous function of s. Hence the limit (3) exists and is equal to the integral (4). We write:

$$(5) \qquad \lim_{n=\infty} \sum_{k=1}^{n} P(x'_k, y'_k) \, \Delta x_k = \int_{(a,\, b)}^{(a',\, b')} P \, dx \qquad \text{or} \qquad \int_C P \, dx.$$

If C is divided into n arcs and the extremities numbered in the inverse order, the new variable (3) approaches as its limit the negative of the former limit. Thus reversing the sense of the integration reverses the sign of the line integral, or

$$(6) \qquad \int_{(a,\, b)}^{(a',\, b')} P \, dx = - \int_{(a',\, b')}^{(a,\, b)} P \, dx,$$

the curve C being the same in both cases.

The limit of (3) is precisely of the type (1), and thus may be written:

$$(7) \qquad \int_C P \cos \tau \, ds.$$

When, however, the sense of the integration is reversed, τ is replaced by $\tau \pm \pi$, and so the sign is changed.

The line integral

$$(8) \qquad \int_{(a,\, b)}^{(a',\, b')} Q \, dy \qquad \text{or} \qquad \int_C Q \, dy$$

is defined in a similar way. Finally,

$$(9) \qquad \lim_{n=\infty} \sum_{k=1}^{n} [P(x'_k, y'_k) \, \Delta x_k + Q(x'_k, y'_k) \, \Delta y_k] = \int_{(a,\, b)}^{(a',\, b')} P \, dx + Q \, dy$$

or

$$\int_C P \, dx + Q \, dy,$$

for this limit is evidently equal to the sum of the line integrals (5) and (8).

The identities (*A*), (*B*), (*C*) are usually referred to in the literature as *Green's Theorem* (1828) or *Gauss's Theorem* (1813). Such identities go back, however, much further, appearing (in the case of volume integrals) as early as 1760/61 in work of Lagrange's. The theorem of Ex. 5 below may, however, properly be called *Green's Theorem*. Cf. a note in the author's *Funktionentheorie*, Vol. I, 2d ed., p. 600.

EXERCISES

1. Extend the integral

$$\int_C \frac{y\,dx - x\,dy}{x^2 + y^2}$$

in the positive sense over the boundary (*i*) of a circle whose centre is at the origin; (*ii*) of a circular ring with its centre at the origin.

Ans. (*i*) -2π; (*ii*) 0.

2. Show that the integral

$$\int_C x\,dy - y\,dx,$$

extended in the positive sense over the complete boundary of any region, is equal to twice the area of the region.

3. Setting $$P = u\frac{\partial v}{\partial x}, \qquad Q = u\frac{\partial v}{\partial y},$$

show that

$$\iint_S \frac{\partial(u, v)}{\partial(x, y)}\,dS = \int_C u\,dv.$$

4. Setting $$P = u\frac{\partial v}{\partial y}, \qquad Q = -u\frac{\partial v}{\partial x},$$

show that

$$\iint_S u\left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2}\right) dS + \iint_S \left(\frac{\partial u}{\partial x}\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\frac{\partial v}{\partial y}\right) dS = -\int_C u\frac{\partial v}{\partial n}\,ds.$$

5. Setting $$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2},$$

show that

$$\iint_S (u\,\Delta v - v\,\Delta u)\,dS = -\int_C \left(u\frac{\partial v}{\partial n} - v\frac{\partial u}{\partial n}\right) ds.$$

This equation is properly known as *Green's Theorem*.

6. Prove that

$$\int_S\!\!\int\left(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}\right)dS = -\int_C \frac{\partial u}{\partial n}\,ds.$$

7. If u is a solution of Laplace's equation,

$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=0,$$

show that

$$\int_S\!\!\int\left[\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial u}{\partial y}\right)^2\right]dS = -\int_C u\frac{\partial u}{\partial n}\,ds.$$

8. If u is a solution of Laplace's equation, show that

$$\int_C \frac{\partial u}{\partial n}\,ds = 0.$$

9. If u is a solution of Laplace's equation, which is not a con stant, show that

$$\int_C u\frac{\partial u}{\partial n}\,ds < 0.$$

5. The Integral $\int_C P\,dx + Q\,dy$.

THEOREM 1. *Let P and Q be two functions which, together with the derivatives $\partial P/\partial y$ and $\partial Q/\partial x$, are continuous within and on the boundary of S. Let*

(1) $$\frac{\partial P}{\partial y}=\frac{\partial Q}{\partial x}$$

at every point of S. Let Σ be any region lying in S; the boundary C of Σ may coincide in part (or wholly) with that of S. Then

$$\int_C P\,dx + Q\,dy = 0,$$

the integral being extended over the complete boundary of Σ in the positive sense.

The proof is given immediately by means of the relation (C) of § 4, since the double integral has the value zero.

THEOREM 2. *Conversely, if P and Q are continuous, together with $\partial P/\partial y$ and $\partial Q/\partial x$, within and on the boundary of S, and if*

$$\int_C P\,dx + Q\,dy = 0,$$

C being the boundary of an arbitrary region Σ lying in S, then

$$\frac{\partial P}{\partial y}=\frac{\partial Q}{\partial x}$$

at every point of S.

Suppose the theorem false. Then the continuous function

$$\frac{\partial P}{\partial y}-\frac{\partial Q}{\partial x}$$

must be different from 0 at some interior point A of S, and hence must be either positive at every point of a suitably chosen neighborhood of A (say, throughout the interior of a small circle about A, lying wholly in S) or else negative throughout such a region. But then the double integral of (C), § 4, could not vanish when extended over this region; and since the line integral which forms the right-hand side of (C) vanishes by hypothesis, we are led to a contradiction. Hence the theorem is established.

EXERCISES

1. Prove by an example that the following theorem is false: Let P and Q be two functions which satisfy the conditions of Theorem 1; and let C be a simple closed curve lying in S. Then

$$\int_C P\,dx + Q\,dy = 0.$$

2. Let S be a ring-shaped region bounded by the curves C_1 and C_2, and let P and Q satisfy the conditions of Theorem 1 in S. Then

$$\int_{C_1} P\,dx + Q\,dy = \int_{C_2} P\,dx + Q\,dy,$$

where each integral is extended in the clock-wise sense over C_1 or C_2.

6. Simply and Multiply Connected Regions. By a *simply connected region* is meant a region such that no closed curve drawn in the region contains in its interior a boundary point of the region. All other regions are called *multiply connected.*

Thus a square or an ellipse is simply connected; more generally, the interior of any simple closed curve, together with the boundary, forms a simply connected region. It is not necessary that the region be finite. The whole plane, or a half-plane, or the region bounded by two rays which emanate from a point, or the (x, y)-plane exclusive of the positive axis of x, — all these are examples of simply connected regions.

A circular ring is an example of a multiply connected region. Consider a region S lying inside a curve C, but outside each of the curves $C_1, \cdots, C_n$. If cuts be made along lines joining the inner boundaries with the outer boundary, the new region, S', will be simply connected. It is clear that n such cuts suffice. These may be drawn in a variety of ways. Thus the curves $C_1, \cdots, C_n$ could be connected in series, and one of them connected also with C. But it can be shown that, no matter how the cuts be drawn, their number will always be the same, namely, n. Such a region is called *doubly* ($n = 1$) or *triply* ($n = 2$) or ($n + 1$)*-tuply* connected.

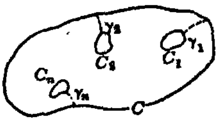


FIG. 54

A simply connected region cannot have a boundary that consists of more than a single piece. But not all regions whose boundary consists of a single piece are simply connected. Thus the exterior of a circle is multiply connected. It is said to be *doubly* connected, since a single cut, as the ray which consists in a radius produced, would yield a simply connected region. Again, the whole plane with the exception of a single point is a doubly connected region.

A simply connected region can also be characterized by the fact that any closed curve drawn in the region can be deformed continuously (like a flexible elastic string) to an interior point of the region — more properly, until it lies wholly within an arbitrarily small neighborhood of the point — without ever coming into collision with the boundary of the region.

Space of Three Dimensions. The ideas and definitions just set forth admit a two-fold generalization to space of three dimensions. Consider the space V between two concentric spheres. In this shell a surface can be drawn (namely, a third concentric sphere) which contains a part of the boundary of V in its interior. Thus we should be led to consider V as multiply connected. But a closed curve drawn in V can be deformed continuously to an interior point of V — *i.e.* until it lies wholly within an arbitrarily small neighborhood of the point — without ever touching the boundary of V. For this reason it is natural to regard V as simply connected. We can meet both situations by saying that V is *linearly simply connected,* but is *multiply connected* with respect to surfaces.

The space (either interior or exterior) bounded by an anchor ring is multiply connected in both senses. If a space is linearly multiply connected, it is obviously multiply connected with respect to surfaces. For, if an arbitrarily slender tube lying in the region could be drawn together continuously to an interior point of the region without colliding with the boundary, the same would be true of a simple closed curve lying within such a tube.

The interior of an anchor ring can be rendered linearly simply connected by introducing a diaphragm, as for example the cut made by a half-plane through the axis. It is not easy to prove that this is true of all the spaces bounded by a finite number of curves and surfaces such as are most familiar to us. So in the following we shall restrict ourselves to spaces that are known to have this property. If n diaphragms are needed to render a given space linearly simply connected, we shall say that the original space was *linearly* $(n+1)$*-tuply connected.* Thus the interior of an anchor ring is linearly doubly connected.

7. The Integral $\int_{(a,b)}^{(x,y)} P\,dx + Q\,dy$.

THEOREM 1. *Let P and Q be continuous, together with $\partial P/\partial y$ and $\partial Q/\partial x$, throughout a region S of the plane. If the integral*

$$(1) \qquad \int_{(a,b)}^{(x,y)} P\,dx + Q\,dy,$$

extended along an arbitrary curve drawn in S, has the same value for all such curves, then

$$(2) \qquad \frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$$

at every point of S.

More generally, the theorem is true if the points (a, b) and (x, y), and the curve joining them, are restricted to lying in a square, the length of whose sides does not exceed a certain positive constant, h, which however may be arbitrarily small, and whose centre may be any point of S.

Let (x_1, y_1) be any interior point of S, which we now hold fast and surround by a square S_1 lying wholly in S. Let C be any simple closed curve lying in S_1, and let (a, b) and (a', b') be two points of C, dividing C into the arcs C_1 and C_2. Since by hypothesis

$$\int_{C_1} P\,dx + Q\,dy = \int_{C_2} P\,dx + Q\,dy,$$

each integral being taken from (a, b) to (a', b'), it is seen that

$$\int_C P\,dx + Q\,dy = 0.$$

Hence by Theorem 2 of § 5 the relation (2) holds throughout S_1 and therefore, in particular, at (x_1, y_1) *. But the latter point was *any* interior point of S. Thus the proposition is proved in all cases.

THEOREM 2. *Let P and Q be continuous, together with $\partial P/\partial y$ and $\partial Q/\partial x$, throughout the interior of a region S of the plane, and let*

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}.$$

If S is simply connected, the integral

$$\int_{(a,b)}^{(x,y)} P\,dx + Q\,dy$$

has the same value for all paths joining (a, b) with (x, y), and thus is a single-valued function u of (x, y). The derivatives of u exist and have the values

$$\frac{\partial u}{\partial x} = P, \qquad \frac{\partial u}{\partial y} = Q.$$

Consider two paths, C_1 and C_2, drawn in S from (a, b) to (x, y). If they meet only at their extremities, they form together a simple closed curve, C, and the integral extended along C has the value 0 by Theorem 1 of § 5. If, however, they meet in other points, a third curve, C_3, can be drawn in S from (a, b) to (x, y) meeting each of the curves C_1 and C_2 only in its extremities, or at most in a finite number of points. The value of the integral for C_3 will be the same as for C_1 or C_2, and hence these latter values will be equal.

FIG. 55

It is seen, then, that the integral defines a single-valued function, u, throughout S. To differentiate u, let (x_0, y_0) be an arbitrary interior point of S. Hold y fast and give to x an increment, Δx. The corresponding increment in u has the value

* We are using here a slight generalization of Theorem 2, which consists in restricting the regions Σ to being *simply connected*. The proof holds good for this more general case. At the time the theorem was stated, simply connected regions had not been introduced.

8. The Integral $\int_{(a,b,c)}^{(x,y,z)} P\,dx + Q\,dy + R\,dz$.

THEOREM 1. *Let P, Q, R, together with the derivatives which enter below, be continuous throughout a linearly simply connected region V of space. In order that the value of the integral*

$$\int_{(a,b,c)}^{(x,y,z)} P\,dx + Q\,dy + R\,dz$$

be the same for all paths joining (a, b, c) and (x, y, z), and lying in V, it is necessary and sufficient that

$$(1) \qquad \frac{\partial Q}{\partial z} = \frac{\partial R}{\partial y}, \qquad \frac{\partial R}{\partial x} = \frac{\partial P}{\partial z}, \qquad \frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}.$$

When these latter conditions are fulfilled, the function u defined by the integral admits derivatives, which are given by the equations:

$$(2) \qquad \frac{\partial u}{\partial x} = P, \qquad \frac{\partial u}{\partial y} = Q, \qquad \frac{\partial u}{\partial z} = R.$$

The proof of the equations (2) is given exactly as in the two-dimensional case, § 7. Moreover, that equations (1) form a necessary condition can be shown by means of the results of § 7, one variable at a time (x or y or z) being held fast. That these equations also represent a sufficient condition will be proved in § 10 by means of Stokes's Theorem. It is, however, possible to give an elementary proof without the aid of Stokes's Theorem; cf. the author's *Funktionentheorie*, Vol. I, Chap. 4, § 3, the method there set forth admitting immediate extension to space of n-dimensions, or to the integral

$$\int_{(a)}^{(x)} P_1\,dx_1 + P_2\,dx_2 + \cdots + P_n\,dx_n.$$

Conditions (1) now take the form:

$$\frac{\partial P_k}{\partial x_l} = \frac{\partial P_l}{\partial x_k}, \qquad k, l = 1, 2, \cdots, n;\ k \neq l$$

For (2) we have

$$\frac{\partial u}{\partial x_k} = P_k, \qquad k = 1, 2, \cdots, n.$$

EXERCISE

Discuss the theorem of the text for the case of a region V which can be rendered linearly simply connected by the introduction of a finite number of diaphragms, like the cuts of the two-dimensional case.

THEOREM 2. *Let P, Q, R be two functions which, together with the derivatives that enter below, are continuous throughout a linearly simply connected region V of space. In order that*

$$\int_C P\,dx + Q\,dy + R\,dz = 0,$$

where C is any simple closed curve lying wholly within V, it is necessary and sufficient that

$$\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} = 0, \qquad \frac{\partial P}{\partial z} - \frac{\partial R}{\partial x} = 0, \qquad \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = 0.$$

The proof of this theorem is also given in § 10.

9. Green's Theorem in Three Dimensions. Let P be a function of (x, y, z), continuous, together with $\partial P/\partial z$, within and on the boundary of a region V. Form the triple integral

$$\iiint_V \frac{\partial P}{\partial z}\,dV$$

It can be evaluated by means of the iterated integral, Chap. IV, § 2:

$$\iiint_V \frac{\partial P}{\partial z}\,dV = \iint_S dS \int_{Z_0}^{Z_1} \frac{\partial P}{\partial z}\,dz$$

$$= \iint_S P(x, y, Z_1)\,dS - \iint_S P(x, y, Z_0)\,dS,$$

where S denotes the projection of V on the (x, y)-plane.

These latter integrals can be expressed in terms of surface integrals taken over the two nappes* of the boundary of V,

* These nappes can be conveniently visualized as follows. Think of V as an opaque solid, and rays of light descending parallel to the axis of z. The part of the boundary illumined will be the upper nappe; the dark part of the boundary, the lower nappe. Moreover, S is the shadow cast by this solid on the (x, y)-plane.

the upper nappe Σ_1, $\quad z = Z_1 = \phi_1(x, y)$;

the lower nappe Σ_0, $\quad z = Z_0 = \phi_0(x, y)$.

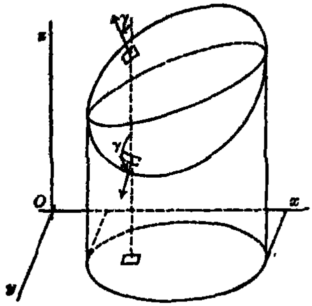


FIG. 56

Let γ denote the angle which the outer normal to the surface Σ of V makes with the positive axis of z. Then

$$\iint_S P(x, y, Z_1)\, dS = \iint_{\Sigma_1} P \cos\gamma\, d\Sigma,$$

$$\iint_S P(x, y, Z_0)\, dS = -\iint_{\Sigma_0} P \cos\gamma\, d\Sigma.$$

Thus the difference of the two double integrals is seen to have the value of the surface integral of $P\cos\gamma$ taken over the entire surface Σ, and so we have the result:

$$\text{(1)} \qquad \iiint_V \frac{\partial P}{\partial z}\, dV = \iint_\Sigma P \cos\gamma\, d\Sigma.$$

Similar formulas could have been obtained if we had started with partial derivatives with respect to x or y, the region V being now projected on the (y, z)-plane or the (x, z)-plane. The results in all three cases can be collected as follows:

$$\text{(2)} \qquad \left\{ \begin{aligned} &\iiint_V \frac{\partial A}{\partial x}\, dV = \iint_S A \cos\alpha\, dS; \\ &\iiint_V \frac{\partial B}{\partial y}\, dV = \iint_S B \cos\beta\, dS; \\ &\iiint_V \frac{\partial C}{\partial z}\, dV = \iint_S C \cos\gamma\, dS, \end{aligned} \right.$$

where we have replaced the letter Σ, as referring to the bounding surface, by the letter S.

If we had used the inner normal, the sign of each right-hand side would have been reversed.

On adding the three equations (2) together, we get:

I. $$\iiint\limits_V \left(\frac{\partial A}{\partial x}+\frac{\partial B}{\partial y}+\frac{\partial C}{\partial z}\right) dV$$

$$= \iint\limits_S (A\cos\alpha + B\cos\beta + C\cos\gamma)\, dS \qquad \text{(outer normal)}$$

$$= -\iint\limits_S (A\cos\alpha + B\cos\beta + C\cos\gamma)\, dS \qquad \text{(inner normal)}.$$

The theorem embodied in this equation (in either form) is of fundamental importance, and we shall point out presently a number of its applications. It will be shown in Chap. XIII that each side of the equation is invariant of any rigid motion of the axes, or of any transformation to other Cartesian axes, provided merely that a right-handed system does not go over into a left-handed system.

In the proofs given or indicated above it is tacitly assumed that the surface of V is cut by a parallel to the axis in question at most in two points or a single line-segment. It is sufficient for the needs of practice to restrict ourselves to such regions V as can be cut up into a finite number of regions V_1, V_2, $\cdots$, for each of which this is true. On writing down equations (2) or I. for each of these regions and adding, the corresponding equation for V results.

These theorems are known in the literature as *Green's Theorem* or *Gauss's Theorem*; cf. § 4, end.

EXERCISES

1. Show that the integral

$$\iint\limits_S (x\cos\alpha + y\cos\beta + z\cos\gamma)\, dS,$$

where α, β, γ refer to the outer normal, is equal to three times the volume of the region.

2. Setting $\quad A = u\dfrac{\partial v}{\partial x}, \quad B = u\dfrac{\partial v}{\partial y}, \quad C = u\dfrac{\partial v}{\partial z},$

show that

$$\iiint\limits_V u\left(\frac{\partial^2 v}{\partial x^2}+\frac{\partial^2 v}{\partial y^2}+\frac{\partial^2 v}{\partial z^2}\right) dV + \iiint\limits_V \left(\frac{\partial u}{\partial x}\frac{\partial v}{\partial x}+\frac{\partial u}{\partial y}\frac{\partial v}{\partial y}+\frac{\partial u}{\partial z}\frac{\partial v}{\partial z}\right) dV$$

$$= -\iint\limits_S u\frac{\partial v}{\partial n}\, dS,$$

where $\partial v/\partial n$ is the directional derivative of v along the inner normal.

3. Setting $$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2},$$
show that
$$\iiint\limits_V (u\Delta v - v\Delta u)\,dV = -\iint\limits_S \left(u\frac{\partial v}{\partial n} - v\frac{\partial u}{\partial n}\right) dS,$$
where n refers to the inner normal. This equation is properly known as *Green's Theorem*; cf. § 4.

4. Prove that
$$\iiint\limits_V \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}\right) dV = -\iint\limits_S \frac{\partial u}{\partial n}\, dS.$$

5. If u is a solution of Laplace's equation,
$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0,$$
show that
$$\iint\limits_S \frac{\partial u}{\partial n}\, dS = 0.$$

6. If u is a solution of Laplace's equation, show that
$$\iiint\limits_V \left[\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial u}{\partial z}\right)^2\right] dV = -\iint\limits_S u\frac{\partial u}{\partial n}\, dS.$$

7. If $u \neq$ const. is a solution of Laplace's equation, show that
$$\iint\limits_S u\frac{\partial u}{\partial n}\, dS < 0.$$

8. Let A, B, C be three functions which, together with the derivatives that enter below, are continuous throughout the interior of a region V' of space. Let V be any region contained within V'; let S refer to the boundary of V, and let α, β, γ be the direction angles of the inner normal of S. In order that
$$\iint\limits_S (A\cos\alpha + B\cos\beta + C\cos\gamma)\, dS = 0,$$
it is necessary and sufficient that
$$\frac{\partial A}{\partial x} + \frac{\partial B}{\partial y} + \frac{\partial C}{\partial z} = 0.$$
Prove this proposition.

10. Stokes's Theorem. Let P, Q, and R be three functions of (x, y, z) which, together with their first partial derivatives, are continuous throughout a region V of space, and let S be a surface lying in V and bounded by the curve C. Let α, β, γ be the *direction angles* of the normal to S, chosen in a suitable sense. Then *Stokes's* Theorem asserts the truth of the equation:

$$\text{I.}\quad \iint_S \left\{\left(\frac{\partial R}{\partial y}-\frac{\partial Q}{\partial z}\right)\cos\alpha+\left(\frac{\partial P}{\partial z}-\frac{\partial R}{\partial x}\right)\cos\beta+\left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)\cos\gamma\right\} dS$$

$$=\int_C P\,dx+Q\,dy+R\,dz,$$

where the line integral is extended over C in a sense dependent on the choice of sense for the normal to S.

The theorem is not true in general for *unilateral* surfaces (cf. *infra*), but it holds for all *two-sided* surfaces. We begin by proving it for a restricted case, and are able then with ease to pass to the general case.

A Restricted Case. Let S be given by the equation

$$(1)\qquad z=\omega(x, y),$$

where ω, together with its first derivatives, is continuous within and on the boundary of a region S' of the (x, y)-plane, and where S lies within the region V. It is furthermore assumed that S' is the kind of region considered in § 4, to which Green's Theorem is applicable.

Consider the integral

$$(2)\qquad \int_C P\,dx+Q\,dy+R\,dz,$$

taken in that sense along C which corresponds to the positive sense of description of the boundary Γ of S'. This integral can be expressed by a line integral over Γ as follows:

$$(3)\qquad \int_\Gamma (P+R\omega_1)\,dx+(Q+R\omega_2)\,dy.$$

For, the value of dz anywhere on S is

$$dz=\omega_1 dx+\omega_2 dy,$$

and along C, dx and dy are the same as along Γ.*

Let $\mathfrak{P}(x, y) = P[x, y, \omega(x, y)] + R[x, y, \omega(x, y)]\,\omega_1(x, y)$,

$\mathfrak{Q}(x, y) = Q[x, y, \omega(x, y)] + R[x, y, \omega(x, y)]\,\omega_2(x, y)$.

Then (3) becomes

(4) $$\int_\Gamma \mathfrak{P}\,dx + \mathfrak{Q}\,dy.$$

This integral can be written in the form (§ 4, C):

(5) $$\iint_{S'} \left(\frac{\partial \mathfrak{Q}}{\partial x} - \frac{\partial \mathfrak{P}}{\partial y}\right) dS'.$$

The integrand of the last integral is seen to have the value:

(6) $$\frac{\partial \mathfrak{Q}}{\partial x} - \frac{\partial \mathfrak{P}}{\partial y} = Q_1 - P_2 + (Q_3 - R_2)\,\omega_1 + (R_1 - P_3)\,\omega_2,$$

where the subscripts against the letters P, Q, R indicate derivatives taken on the hypothesis that (x, y, z) are the independent variables.

Let the positive sense of the normal to S be defined as that for which the direction angle γ is acute. Then

$$\cos\alpha = -\omega_1/\Delta, \quad \cos\beta = -\omega_2/\Delta, \quad \cos\gamma = 1/\Delta, \quad \Delta = \sqrt{1 + \omega_1^2 + \omega_2^2}.$$

Hence the integral (5) can, by the aid of (6), be written in the form

(7) $$\iint_{S'} \{(Q_1 - P_2) + (R_2 - Q_3)\,\Delta\cos\alpha + (P_3 - R_1)\,\Delta\cos\beta\}\,dS'.$$

* A fuller explanation of this point is as follows. Let C be given in the parametric form:

C: $\quad x = f(\lambda), \quad y = \phi(\lambda) \quad z = \psi(\lambda), \qquad 0 \leqq \lambda \leqq 1.$

Then the integral (1) becomes

$$\int_0^1 (Px' + Qy' + Rz')\,d\lambda.$$

But from (1)

$$z' = \omega_1 x' + \omega_2 y'.$$

Hence this integral has the value

$$\int_0^1 \{(P + R\omega_1)\,x' + (Q + R\omega_2)\,y'\}\,d\lambda.$$

On the other hand, the curve Γ is represented by the equations

Γ: $\quad x = f(\lambda), \quad y = \phi(\lambda), \qquad 0 \leqq \lambda \leqq 1$

Hence the last integral is the same as the integral (3).

This integral has the same value as the surface integral, taken over S:

$$(8)\quad \iint_S \left\{\left(\frac{\partial R}{\partial y}-\frac{\partial Q}{\partial z}\right)\cos\alpha+\left(\frac{\partial P}{\partial z}-\frac{\partial R}{\partial x}\right)\cos\beta+\left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)\cos\gamma\right\} dS.$$

But this is precisely the integral that stands on the left of Equation I., and hence the theorem is proved for this case. It is to be observed that the positive sense of C, the inner normal to C regarded as the bounding curve of S, and the positive normal of S are so oriented to each other as the positive axes of x, y, and z respectively.

An Invariant Property. It will be shown in Chap. XIII that, if the coordinates are transformed to any new system of Cartesian axes, provided merely that a right-handed system does not go into a left-handed system, the integrands of both the line integral and the surface integral in I. will preserve their form. Thus

$$P'\frac{dx'}{d\lambda}+Q'\frac{dy'}{d\lambda}+R'\frac{dz'}{d\lambda}=P\frac{dx}{d\lambda}+Q\frac{dy}{d\lambda}+R\frac{dz}{d\lambda};$$

and similarly, the integrand of the transformed surface integral will be *

$$\left(\frac{\partial R'}{\partial y'}-\frac{\partial Q'}{\partial z'}\right)\cos\alpha'+\left(\frac{\partial P'}{\partial z'}-\frac{\partial R'}{\partial x'}\right)\cos\beta'+\left(\frac{\partial Q'}{\partial x'}-\frac{\partial P'}{\partial y'}\right)\cos\gamma'.$$

But Equation I. is invariant even of a reflection, as $z'=-z$, if $R'=-R$; for then $\gamma'=\pi-\gamma$ and the sense of C is reversed.

Suppose now that we have an arbitrary bounded surface, S. Then we can cut it up into a finite number of pieces, $S_1, S_2, \cdots$, each of which, referred to a system of Cartesian axes properly chosen, will come under the case just treated. Hence Stokes's theorem will hold for such a piece, no matter how the axes are chosen, and so we may refer all the pieces to the same axes.

Write down, then, Stokes's theorem for each of the pieces $S_1, S_2, \cdots$, and add the results together. For the kind of surfaces we most readily think of, like a piece of a sphere or a paraboloid, we shall be integrating along each of the cuts once in one direction and once in the opposite direction. So these contributions to the sum on the

* These facts can, however, be proved here directly by the student, by merely writing down the most general equations which represent such a transformation (*Analytic Geometry*, p. 592 and p. 594), and then computing the original integrands in terms of the new variables.

right-hand side will annul one another, and we shall have remaining merely the line integral over the complete boundary of S, taken in the proper sense.

FIG. 57

This sense will depend on which normal to S we have chosen as positive. Having made this choice at a single point, we can determine it at all other points by the method of continuity. Think of the surface as a material surface, and think of a thumb tack as placed against it, the shaft pointing in the right direction at a single point. Allow the thumb tack to slide about on the surface. Then the proper sense for the normal will be uniquely determined at all other points. Such a surface is called *bi-lateral*, as having *two sides*.

Unilateral Surfaces. But there are surfaces which do not have this property; as was first shown by Möbius. Take a rectangular strip of paper and bring the ends together, allowing A and B to fall, not on C and D, but on D and C respectively. Then the thumb tack can be slid on the surface, — say, along the long central line, — so as to come back to the starting point reversed in sense. This makes trouble for the direction angles α, β, γ of the surface integral.

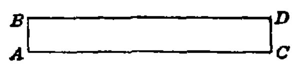


FIG. 58

On the other hand, although this surface can readily be cut up into pieces S_1, S_2, $\cdots$ of the kind desired and a positive sense for the boundary be chosen for one of them, positive senses cannot be assigned to the others so that the integrations along the cuts will all cancel. Try it.

Such a surface is called *unilateral*, for it has but one side. If a painter agreed to paint only one side of the surface, the Union would interfere.

The Final Condition in Stokes's Theorem. We are now in a position to complete the statement of Stokes's Theorem. It is, that the surface in question be *bi-lateral*. Then the proof goes through as set forth above.

An Application. Let V be an arbitrary linearly simply connected region of space, and let C be any simple closed curve drawn in V. Then C can be drawn together continuously, always remaining simple and wholly within V, into a curve C_1 lying in a sphere K contained in V. Let P, Q, and R be three functions which, together with the derivatives that enter below, are continuous throughout the whole interior of V, and let

$$\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} = 0, \qquad \frac{\partial P}{\partial z} - \frac{\partial R}{\partial x} = 0, \qquad \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = 0.$$

Consider the integral

$$\int_{C} P\,dx + Q\,dy + R\,dz.$$

Its value will be the same for all intermediate positions of C. For, two near positions form the complete boundary of a bi-lateral surface contained in V, and hence this integral, extended over both positions of C (but in opposite senses) vanishes by Stokes's Theorem.

Let M be a diametral plane of the sphere K. Then C_1 can be deformed continuously toward its projection on M. Again, the value of the integral remains constant. The limiting position of C_1 is, however, a closed curve Γ (no longer simple, in general) which lies in M. But for a closed *plane* curve, simple or not, the integral vanishes by § 5. Hence the original integral $= 0$.

Finally, the integral

$$\int_{(a,\,b,\,c)}^{(x,\,y,\,z)} P\,dx + Q\,dy + R\,dz$$

has the same value for all paths connecting (a, b, c) with (x, y, z) and lying in V, and hence it defines a single-valued function, u, in V.

Thus all the theorems of § 8 are established. Stokes's Theorem owes its importance, however, chiefly to those cases in physics, in which the surface integral has a meaning.

11. Flow of Heat. Imagine a slab of copper 2 cm. thick, with one side packed in melting ice at temperature $u = u_0 = 0°$, and the other side exposed to steam, $u = u_1 = 100°$. A flow of heat within the slab results, and if the above surface temperatures are permanently maintained, the flow will tend toward a limiting condition, in which the lines of flow are the perpendiculars to the faces of the slab, and the isothermal surfaces are the planes parallel to these faces. Moreover, the temperature will fall off steadily, as a point P traces a line of flow. If x denotes the distance of P from the surface of temperature u_1, and a, the thickness of the plate, then

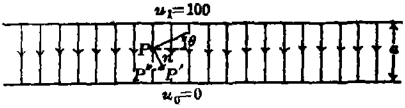


FIG. 59

(1) $$u = u_1 - (u_1 - u_0)\frac{x}{a}.$$

All this is plausible enough, but how do we *know* it is so, and what do we mean by heat, anyway? To answer the second question first, we think of heat as an imponderable substance which can flow freely in a conductor and which can be measured in calories,* as sugar in pounds. And now the above statements about heat, including equation (1) above and equation (2) below, are no more and no less than *physical laws*, — the facts of nature we take for granted.

To go on: Consider a plane region S, of area A, situated in the slab and parallel to the faces. Let Q be the quantity of heat which traverses this surface in one second. Then obviously† Q is proportional to A, to the difference in temperature of the faces, and inversely to the thickness of the plate:

$$Q \propto A, \quad u_1 - u_0, \quad \frac{1}{a};$$

(2)
$$Q = K\frac{u_1 - u_0}{a}A,$$

where K is a physical constant, the *specific conductivity*.

Next, let the plane area S be oblique to the faces, making an angle θ with them. Then the amount of heat, Q, which traverses the surface in one second will obviously be the same as that which traverses the projection, $A\cos\theta$, of S on a face, or

(3)
$$Q = K\frac{u_1 - u_0}{a}A\cos\theta.$$

The Normal Derivative. Consider the normal drawn to S in the sense of the flow. Let n be its length. Then it is readily seen that

(4)
$$\frac{\partial u}{\partial n} = \frac{\partial u}{\partial x}\cos\theta.$$

For
$$u_{p'} = u_{p''}, \qquad \Delta x = \Delta n\cos\theta,$$

$$\frac{\Delta u}{\Delta n\cos\theta} = \frac{\Delta u}{\Delta x}, \qquad \Delta u = u_{p'} - u_p = u_{p''} - u_p,$$

and it remains merely to take the limits.

* By a *calorie* is meant the quantity of heat required to raise one gramme of water one degree centigrade (the initial temperature being 15°).

† Each statement is a physical law. "Obviously" means merely that these laws are easily accepted.

From (1) we find:

$$\frac{\partial u}{\partial x} = -\frac{u_1 - u_0}{a}.$$

On combining this equation with (3) and eliminating $\partial u/\partial x$ by means of (4), we obtain:

$$Q = -K\frac{\partial u}{\partial n}A. \tag{5}$$

This result embodies all the physical laws that have gone before, for from it we can deduce both (1) and (3). Moreover, it states these laws in terms of what is going on in the neighborhood of P, and not in terms of the temperature at remote points.

If the sense of the normal be reversed, Q will be replaced by its negative.

12. Continuation. The General Case. Consider now an arbitrary steady flow. The lines of flow will be curved lines, forming a two-parameter family of space curves which just fill out the region of flow. These curves are obtained by considering the family of isothermal surfaces,

$$u = u_0. \tag{6}$$

In the neighborhood of any point P within the region of flow, the situation is similar to that set forth in § 11, for the portions of these surfaces contained in the neighborhood of P look almost like planes, which are sensibly parallel to one another, and so the lines of flow are seen to be curves cutting these surfaces orthogonally.

Let S be a surface, open or closed, which lies in the region of flow. Cut S up into n pieces $\Delta S_1, \cdots, \Delta S_n$, the maximum diameters of these pieces being small. Then each of the pieces will look like a small piece of a plane surface, and it is physically evident that the amount of heat which traverses the k-th region in one second will be approximately

$$-K\left(\frac{\partial u}{\partial n}\right)_k \Delta S_k,$$

where the normal derivative is formed at an arbitrary point of that region. Thus the quantity, Q, of heat which traverses the whole surface S in one second will be approximately

$$\sum_{k=1}^{n} -K\left(\frac{\partial u}{\partial n}\right)_k \Delta S_k,$$

and the approximation will be closer and closer, the smaller the subregions, or

$$Q = \lim_{n=\infty} \sum_{k=1}^{n} - K\left(\frac{\partial u}{\partial n}\right)_k \Delta S_k = -\iint_S K \frac{\partial u}{\partial n} ds.$$

This whole deduction has been heuristic. We have dwelt on the physical pictures and considered what it is reasonable to expect. To justify in this manner the final result, it would be necessary to make the physical assumptions sharper, in order to draw mathematical inferences from them. After all that has been done, the final result is the equation

$$Q = -\iint_S K \frac{\partial u}{\partial n} dS. \tag{7}$$

It is, therefore, simpler, both physically and mathematically, now that we see what it is reasonable to expect, to begin at this end and *lay down equation* (7) *as the one physical law.*

Unsteady Flow. In the case of an arbitrary flow an instantaneous photograph of the lines of flow at one instant would be different from that at another instant.* Nevertheless, these lines do not shift abruptly, and for a short interval of time succeeding an arbitrary instant the rate of flow across S is given approximately by (7), or

$$\frac{\Delta Q}{\Delta t} = -\iint_S K \frac{\partial u}{\partial n} dS + \zeta,$$

where ζ approaches 0 with Δt. Thus we have

$$\lim_{\Delta t \doteq 0} \frac{\Delta Q}{\Delta t} = \frac{\partial Q}{\partial t} = -\iint_S K \frac{\partial u}{\partial n} dS.$$

Again, the reasoning has been heuristic. We have been shooting in the target; and now we can bring all the physical assumptions considered from the beginning into the one

Physical Hypothesis. *In the case of an arbitrary flow of heat, the rate at which heat traverses a fixed surface S is*

$$\frac{\partial Q}{\partial t} = -\iint_S K \frac{\partial u}{\partial n} dS. \tag{8}$$

* Whereas in the case of a steady flow the lines of flow formed a *two-parameter* family of curves, the lines of flow in the general case correspond to the individual particles of the substance at an initial instant, and thus form a *three-parameter* family.

We have treated K as constant, and it is so for ordinary substances and moderate variations in temperature. The Hypothesis includes the case, however, that K is a continuous function of x, y, z, t.

EXERCISE

Show that the lines of flow are given by the simultaneous system of differential equations:

$$\frac{dx}{\frac{\partial u}{\partial x}} = \frac{dy}{\frac{\partial u}{\partial y}} = \frac{dz}{\frac{\partial u}{\partial z}}.$$

13. A New Heat Problem. If a homogeneous substance be raised from the constant initial temperature u_0 to the constant final temperature u_1, the quantity of heat required, Q, will be proportional to the rise in temperature and the volume:

$$Q \quad \propto \quad u_1 - u_0, \quad V;$$

(9) $$Q = C(u_1 - u_0)V,$$

where C is a physical constant depending on the substance, the *specific heat per unit of volume.*

If, now, an arbitrary homogeneous substance be raised from the continuous initial temperature u_0 to the continuous final temperature u_1, the amount of heat required will be

(10) $$Q = \iiint\limits_V C(u_1 - u_0)\,dV,$$

as is shown by the usual procedure of the integral calculus.

Consider an arbitrary flow of heat. Let the temperature,

$$u = f(x, y, z, t),$$

be u_0 when $t = t_0$ and u_1 when $t = t_0 + \Delta t$. The quantity of heat required to produce this change is given by (10), where for Q we now write ΔQ.

On the other hand,

$$u_1 - u_0 = \Delta t f_t(x, y, z, t_0 + \theta\Delta t), \qquad 0 < \theta < 1.$$

Hence

$$\frac{\Delta Q}{\Delta t} = \iiint\limits_V C f_t(x, y, z, t_0 + \theta\Delta t)\,dV.$$

When Δt approaches 0, the integral approaches the integral of the limiting function, as will be shown under the proof of Leibniz's Rule. Thus

$$\lim_{\Delta t \doteq 0} \frac{\Delta Q}{\Delta t} = \frac{\partial Q}{\partial t} = \iiint_V C \frac{\partial u}{\partial t} dV.$$

We can, however, assemble all our physical hypotheses into the single one:

Physical Hypothesis. *If in the case of a flow of heat the temperature is a continuous function of x, y, z, t, and if this is true of $\partial u/\partial t$, too, then the rate at which the heat is accumulating in a given region, V, is given by the formula:*

$$(11) \qquad \frac{\partial Q}{\partial t} = \iiint_V C \frac{\partial u}{\partial t} dV,$$

where the specific heat, C, is either a constant or a continuous function of x, y, z, t.

From (11) equations (9) and (10) follow at once.

14. The Heat Equation. Consider an arbitrary flow of heat, in which the temperature, together with the partial derivatives* of the first two orders, is continuous in x, y, z, t. Let V be an arbitrary sub-region contained in the region of flow. Then the rate at which the heat in V is increasing is given in two forms, namely, by equation (8) of § 12 and by (11) of § 13. Hence

$$(12) \qquad \iiint_V C \frac{\partial u}{\partial t} dV = -\iint_S K \frac{\partial u}{\partial n} dS,$$

where n refers to the *inner normal* of S. For ordinary substances and moderate variations in the temperature, K may be assumed constant.

By Green's Theorem, § 9, we have:

$$\iiint_V \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}\right) dV = -\iint_S \frac{\partial u}{\partial n} dS$$

Hence

$$\iiint_V \left[C\frac{\partial u}{\partial t} - K\Delta u\right] dV = 0.$$

* It is not necessary to extend this requirement to *all* these derivatives; but the loss in generality is unimportant.

The integrand of this last integral is continuous throughout the whole region of flow. Hence it must vanish at every point of the region:

$$C\frac{\partial u}{\partial t} - K\Delta u = 0.$$

For, if it were, for example, positive at a point P within the region, it would be positive throughout a certain neighborhood of P. But the integral of a positive function cannot be zero.

We thus arrive at the differential equation which governs the flow of heat in the general case:

$$(13) \qquad \frac{\partial u}{\partial t} = a^2\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}\right), \qquad a^2 = \frac{K}{C}.$$

It is a linear partial differential equation of the second order with constant coefficients.

Steady Flow. We can now define a *steady flow* as one in which the temperature at any given point is independent of the time. From (13) it follows that the temperature, in the case of a steady flow, will satisfy Laplace's Equation:

$$(14) \qquad \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0.$$

Conversely, if the temperature satisfies Laplace's Equation, then from (13) $\partial u/\partial t = 0$, and the flow is steady.

A necessary and sufficient condition for a steady flow is the following: if V be an arbitrary sub-region contained in the region of flow, then $\partial Q/\partial t = 0$ for this region.

The latter property might be taken as the definition of a steady flow.

EXERCISE

If K is variable, but continuous, together with its partial derivatives of the first order, show that the heat equation becomes:

$$\frac{\partial u}{\partial t} = \frac{1}{C}\left\{\frac{\partial}{\partial x}\left(K\frac{\partial u}{\partial x}\right) + \frac{\partial}{\partial y}\left(K\frac{\partial u}{\partial y}\right) + \frac{\partial}{\partial z}\left(K\frac{\partial u}{\partial z}\right)\right\}.$$

Here, C may also be variable; it will be continuous.

15. Flow of Electricity in Conductors. The flow of electricity in a conductor is mathematically identical with the problem of the flow of heat just discussed. On replacing throughout the word *heat* by *electricity*, and the word *temperature* by *potential*, the foregoing treatment applies to the electrical case.

16. Two-Dimensional Flow. Consider a solid cylinder of arbitrary cross-section, S, cut off by two planes perpendicular to its elements. Let the ends be insulated, and let the remaining surface be maintained at prescribed temperatures which shall not change with the time. In particular, the surface temperature shall be a continuous function, and it shall be constant along each element of the cylinder.

The limiting flow, *i.e.* the steady flow which corresponds to the surface conditions, will be one in which the lines of flow all lie in planes parallel to the bases; and the lines of flow in one of these planes project on the lines of flow in any other plane.

Thus the flow is completely described by the flow in one of these planes. Let the (x, y)-axes be chosen in this plane. Then u does not depend on z; hence $\partial^2 u/\partial z^2 = 0$, and Laplace's Equation reduces to

$$(1) \qquad \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0.$$

Equation (8) of § 12 reduces to

$$(2) \qquad \frac{\partial Q}{\partial t} = -\int_C K \frac{\partial u}{\partial n} ds,$$

provided that the altitude of the cylinder is unity.

A necessary and sufficient condition for a steady flow, when K is constant, is that

$$(3) \qquad \int_\Gamma \frac{\partial u}{\partial n} ds = 0$$

for every sub-region, the integral being extended over the complete boundary in the positive sense. Equation (3) follows from Laplace's Equation, and conversely; § 4.

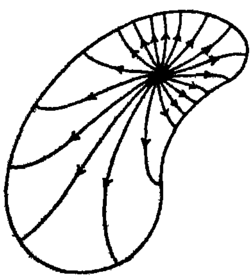


Fig. 60

Flow of Electricity. A two-dimensional flow can be realized as follows. Consider a piece of tin foil. Let the edge be connected with a thick piece of copper, and let one pole of a battery be connected with the copper; the other, with an interior point of the tin foil. Then a flow of electricity in the tin foil will be established, and since the resistance of the copper is negligible, while that of the tin foil is not, the edge of the tin foil will be at constant potential.*

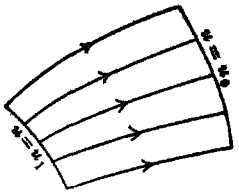


FIG. 61

Again, two segments of the boundary of a piece of tin foil can be connected with two thick pieces of copper, and these in turn with the poles of a battery. A flow of electricity is thus set up in the tin foil, in which the other two edges of the latter are lines of flow.

17. Continuation. The Conjugate Function. Equation (3) of § 16 can be written in the equivalent form:

$$\int_{\Gamma} -\frac{\partial u}{\partial y}\,dx + \frac{\partial u}{\partial x}\,dy = 0. \tag{4}$$

This form suggests that we consider the function

$$v = \int_{(a,b)}^{(x,y)} -\frac{\partial u}{\partial y}\,dx + \frac{\partial u}{\partial x}\,dy + C, \tag{5}$$

and that we take the region S as simply connected. This integral is independent of the path of integration, as follows from (4), or from the direct application of the test of § 7:

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}, \qquad -\frac{\partial^2 u}{\partial y^2} = \frac{\partial^2 u}{\partial x^2},$$

the latter condition being fulfilled since u is harmonic.

From (5) we infer the following relations:

$$\begin{cases} \dfrac{\partial u}{\partial x} = \dfrac{\partial v}{\partial y}, \\[2ex] \dfrac{\partial u}{\partial y} = -\dfrac{\partial v}{\partial x}. \end{cases} \tag{6}$$

A pair of functions, u and v, which are continuous, together with their first partial derivatives, and which satisfy equations (6), are said to be *conjugate*. More precisely, v is conjugate to u, and $-u$ is conjugate to v.

* The same flow could be realized by taking the region S as a non-conducting surface and covering it on both sides and along the edge with tin foil. On connecting the two poles of a battery with points of the tin foil above one another, the flow in question ensues.

The curves $v = \text{const.}$ are seen to cut the curves $u = \text{const.}$ orthogonally. For,

$$\frac{\dfrac{\partial u}{\partial x}}{\dfrac{\partial u}{\partial y}} = -\left[\frac{\dfrac{\partial v}{\partial x}}{\dfrac{\partial v}{\partial y}}\right]^{-1}.$$

Thus the curves of the former family are the lines of flow; for, they cut the isothermals at right angles.

EXERCISES

1. If u and v are conjugate functions, and if they possess continuous partial derivatives of the second order,* show that both u and v are harmonic functions.

2. Show that, if u and v are conjugate functions, they satisfy the differential equations in polar coordinates:

$$\frac{\partial u}{\partial r} = \frac{1}{r}\frac{\partial v}{\partial \theta}, \qquad \frac{1}{r}\frac{\partial u}{\partial \theta} = -\frac{\partial v}{\partial r}.$$

3. Prove that Laplace's Equation becomes, in polar coordinates:

$$r^2\frac{\partial^2 u}{\partial r^2} + r\frac{\partial u}{\partial r} + \frac{\partial^2 u}{\partial \theta^2} = 0.$$

The equation may be written suggestively in the form:

$$\frac{\partial}{\partial \log r}\left(\frac{\partial u}{\partial \log r}\right) + \frac{\partial^2 u}{\partial \theta^2} = 0.$$

18. Theory of Functions of a Complex Variable. Equations (6) of § 17 are those which express the condition that a function of a complex variable:

(1) $$w = f(z),$$

$$z = x + yi, \qquad w = u + vi, \qquad i = \sqrt{-1},$$

should possess a derivative; *cf.* Chap. XX, § 11:

(2) $$D_z w = f'(z).$$

Such a function, $f(z)$, is said to be *analytic*. Since its real part, u, satisfies Laplace's Equation, we have an unlimited source of harmonic functions. Thus

* As a matter of fact, this condition is satisfied automatically; but this is not the place to prove that theorem.

(3) $$w = z^2$$

is seen to have a derivative, and hence

(4) $$u = x^2 - y^2, \qquad v = 2xy$$

form a pair of conjugate functions.

EXERCISES

1. Show that any harmonic function which depends on r alone, but not on θ, is of the form:

$$u = c \log r + c'.$$

2. The isothermals of the function u of Question 1 are concentric circles. Show that, when the particular values $u_1, u_2, \cdots, u_n$ form an arithmetic progression, the corresponding values of r form a geometric progression.

3. Discuss the conjugate family.

4. Draw the curves $u =$ const. and $v =$ const. in the first quadrant, corresponding to the functions (4) above.

5. Describe two cases of flow of electricity corresponding to the results of Question 4, and show precisely how to realize the flow in each case.

19. Irrotational Flow of an Incompressible Fluid. The domain of ideas which we have just been considering — the physical pictures and the mathematical treatment — is closely related to that of a flow of a fluid in two dimensions. If the density, ρ, is constant as regards both space and time, the equation of continuity, Chap. XII, § 9, reduces to

(1) $$\frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} = 0,$$

where X and Y are the components of the velocity at any point along the axes.

The condition that the flow be *irrotational* is that a velocity potential, u, exist:

(2) $$X = \frac{\partial u}{\partial x}, \qquad Y = \frac{\partial u}{\partial y},$$

and if the flow is to be steady, these derivatives must be independent of the time.

From (1) and (2) it follows that u is harmonic; and conversely any harmonic function u, independent of t, leads through (2) to an irrotational, steady flow of an imponderable fluid of constant density.

References. The formal transformations based on Green's Theorem are treated in much detail in the opening chapter of Watson and Burbury's *Electricity and Magnetism.* The flow of heat is well set forth in Fourier's *Analytic Theory of Heat,* translated by Freeman. See also the references at the close of Chap. XII and Chap. XX.

CHAPTER XII

TRANSFORMATION OF MULTIPLE INTEGRALS. EQUATION OF CONTINUITY

The definite integral of a function of a single variable was defined as the limit of a sum, and the existence of this limit was based on the geometric evidence of the area under a curve. It was possible to extend the method to double integrals; but for triple integrals geometric intuition broke down, since a four-dimensional space would be needed.

By means of the new formulation of the arithmetic definition of the definite integral, to which we now turn, the above gap relating to triple integrals is readily filled, and, on the other hand, the theorems concerning iterated integrals and change of variables admit simple treatment.

1. A New Definition of the Definite Integral. Let $f(x)$ be a continuous function of x in the interval

$$a \leqq x \leqq b.$$

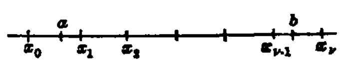


FIG. 62

Let the axis of x be divided up into segments, each of length $1/2^n$, by means of the points $x = l/2^n$, where l is any whole number, — positive, negative, or zero. Denote those points which lie actually within the interval (a, b) by $x_1, \cdots, x_{\nu-1}$ (where $x_{k-1} < x_k$), and let x_0, x_ν be the end-points, or the next points outside the interval:

$$x_0 \leqq a < x_1, \qquad x_{\nu-1} < b \leqq x_\nu.$$

Next, form the sum

$$S_n = \sum_{k=p}^{q} f(x'_k)\,\Delta x_k,$$

where x'_k is any point which lies in the interval $x_{k-1} \leqq x \leqq x_k$ and in which $f(x)$ is defined, and

$$\Delta x_k = x_k - x_{k-1}; \qquad p = 1 \text{ or } 2; \qquad q = \nu - 1 \text{ or } \nu,$$

i.e. p may be chosen at pleasure to be either of the numbers 1 or 2; and q, independently of p, to be either of the numbers $\nu - 1$ or ν.

We proceed to show that S_n approaches a limit when n increases without limit, and we shall define this limit to be the *definite integral* of the function $f(x)$ from a to b:

$$\lim_{n=\infty} S_n = \int_a^b f(x)\,dx.$$

Let M_k, m_k be respectively the largest and the smallest values of the function $f(x)$ in the interval $x_{k-1} \leqq x \leqq x_k$ (or in that part of the interval in which $f(x)$ is defined). Then

$$m_k \leqq f(x'_k) \leqq M_k$$

and
$$\sum_{k=2}^{\nu-1} m_k\,\Delta x_k \leqq \sum_{k=2}^{\nu-1} f(x'_k)\,\Delta x_k \leqq \sum_{k=2}^{\nu-1} M_k\,\Delta x_k.$$

And now the proof consists in showing that each of the extreme sums in the double inequality approaches a limit, and that these limits are equal. We will restrict ourselves for the present to the case that $f(x) > 0$.

Consider the sum

(1)
$$\sum_{k=1}^{\nu} M_k\,\Delta x_k.$$

When n increases by 1, each interval (x_{k-1}, x_k) is bisected, and the M's corresponding to the two halves are at most equal to the former M_k. The modification in statement for the extreme intervals, (x_0, x_1) and $(x_{\nu-1}, x_\nu)$, is obvious.

Hence the value of the sum (1), if it changes, decreases. But it will never be less than

$$m(b-a),$$

where m is the minimum value of $f(x)$ in the interval (a, b).

It follows, then, that the sum (1) approaches a limit; *Introduction to the Calculus*, p. 391. Denote the value of the latter by A. Then, obviously, we also have

$$\lim_{n=\infty} \sum_{k=2}^{\nu-1} M_k\,\Delta x_k = A.$$

Next, form the sum

(2)
$$\sum_{k=2}^{\nu-1} m_k\,\Delta x_k.$$

Analogous considerations show that this sum likewise approaches a limit,* which we will denote by B:

$$\lim_{n=\infty} \sum_{k=2}^{\nu-1} m_k \Delta x_k = B.$$

Since

$$\sum_{k=2}^{\nu-1} M_k \Delta x_k - \sum_{k=2}^{\nu-1} m_k \Delta x_k = \sum_{k=2}^{\nu-1} (M_k - m_k) \Delta x_k \geqq 0,$$

it follows that $A \geqq B$. We wish to show that only the lower sign can hold.

To prove this, it is sufficient to show that, when a positive quantity ϵ has been chosen at pleasure, it is then possible to find an n: $n = N$, such that

(3) $$M_k - m_k < \epsilon$$

for all values of k in question. Obviously, relation (3) will then hold for all larger values of n. The proof of this theorem depends on the property of *uniform continuity* — a subject belonging to a more advanced stage of analysis; *cf.* for example the author's *Funktionentheorie*, Vol. I, Chap. I, § 4.

Thus the variables (1) and (2) approach one and the same limit, A, and hence S_n approaches this limit, too, no matter how x_k' is chosen in the interval (x_{k-1}, x_k). We define this limit as the *definite integral* of $f(x)$ and write:

$$\lim_{n=\infty} \sum_{k=p}^{q} f(x_k') \Delta x_k = \int_a^b f(x)\, dx.$$

Corollary 1. LAW OF THE MEAN:

$$\int_a^b f(x)\, dx = f(\xi)\,(b - a), \qquad a < \xi < b.$$

Here, ξ denotes a point properly chosen within the interval (a, b).

* This result can, however, be deduced immediately from the theorem just proved. For, let a constant, C, be chosen greater than the greatest value of $f(x)$ in the interval (a, b). Form the function

$$F(x) = C - f(x).$$

Then the sum (1), formed for $F(x)$, has the value

$$C(x_\nu - x_0) - \sum_{k=1}^{\nu} m_k \Delta x_k.$$

Since this whole variable approaches a limit, and since its first term also approaches a limit, its last term must likewise converge.

Corollary 2. If c be any intermediate point, $a < c < b$, then

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx.$$

The proof of each corollary follows immediately from the definition of the integral.

The Most General Law of Sub-Division. Let the points $x_0, x_1, \cdots, x_\nu$, now be chosen arbitrarily, subject merely to the conditions that

$$x_0 \leqq a < x_1, \quad x_{k-1} < x_k, \quad x_{\nu-1} < b \leqq x_\nu, \quad \lim_{n=\infty} \Delta x_{k_1} = 0,$$

where Δx_{k_1} is the maximum Δx_k for the n-th sub-division. Let

$$\overline{S}_n = \sum_{k=p}^{q} f(x'_k)\,\Delta x_k = \sum_{k=2}^{\nu-1} f(x'_k)\,\Delta x_k + e_1 f(x'_1)\,\Delta x_1 + e_2 f(x'_\nu)\,\Delta x_\nu,$$

$$e_1, e_2 = 0, 1.$$

On the other hand,

$$\int_{x_1}^{x_{\nu-1}} f(x)\,dx = \sum_{k=2}^{\nu-1} \int_{x_{k-1}}^{x_k} f(x)\,dx = \sum_{k=2}^{\nu-1} f(x''_k)\,\Delta x_k, \qquad x_{k-1} \leqq x''_k \leqq x_k.$$

Hence *

$$\overline{S}_n - \int_a^b f(x)\,dx = \sum_{k=2}^{\nu-1} [f(x'_k) - f(x''_k)]\,\Delta x_k$$

$$+ e_1 f(x'_1)\,\Delta x_1 + e_2 f(x'_\nu)\,\Delta x_\nu - \int_a^{x_1} f(x)\,dx - \int_{x_{\nu-1}}^b f(x)\,dx.$$

When n increases without limit, each term on the right approaches zero as its limit, and hence $\overline{S}_n$ approaches a limit. The value of this limit is

$$\int_a^b f(x)\,dx.$$

* To be precise, we need here the further definition:

$$\int_c^c f(x)\,dx = 0,$$

where c is any point of the interval (a, b).

It remains merely to remove the restriction that $f(x) > 0$. If this condition is not fulfilled, form the function

$$\phi(x) = f(x) + C,$$

where C is so chosen that $\phi(x) > 0$. Since $\phi(x)$ possesses an integral, it can now be shown without difficulty that $f(x)$ also possesses an integral; that Corollaries 1, 2 hold for this integral, and that the general sum $\overline{S}_n$ approaches this integral as its limit.

2. Continuation. Multiple Integrals. The foregoing definition has the advantage that it admits immediate extension to multiple integrals.

Double Integrals. Let S be a finite region of the (x,y)-plane, whose boundary is made up of a finite number of arcs, each of which can be represented in at least one of the forms

$$(1) \qquad y = \phi(x), \qquad x = \omega(y),$$

where the function standing on the right-hand side is continuous throughout a certain interval,

$$a \leqq x \leqq b, \qquad \text{or} \qquad \alpha \leqq y \leqq \beta.$$

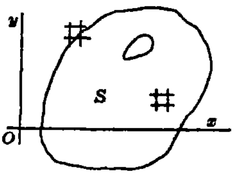


FIG. 63

Let $f(x,y)$ be a function continuous within and on the boundary of S.

We next divide the plane into equal squares by the lines

$$x = \frac{i}{2^n}, \qquad y = \frac{j}{2^n},$$

where i and j, independently of each other, range through all integral values.

Consider (a) those squares which lie wholly within S; and (b) those squares which contain, in their interior or on their boundary, at least one boundary point of S. Let (x_k, y_k) be an arbitrary point of S lying within or on the boundary of one of these squares. Form the sum

$$S_n = \sum f(x_k, y_k)\,\Delta S_k,$$

where ΔS_k denotes the area of the square (and is the same for all squares, namely 2^{-2n}) and the summation *must* include all the squares of Class (a) and *may* include all, some, or none of the squares of Class (b). *Then S_n approaches a limit when n becomes infinite.* The value of this limit is defined as the value of the *double integral* of $f(x, y)$ extended over the region S:

$$\lim_{n=\infty} S_n = \iint_S f(x, y)\, dS.$$

To prove the theorem, let M_k, m_k denote respectively the largest and the smallest value of $f(x, y)$ in a given square. Form the sums

$$\sum{}^{+} M_k \Delta S_k, \qquad \sum{}^{-} m_k \Delta S_k,$$

where the first sum is extended to all the squares of (a) and (b), and the second, to the squares of (a) only.

The further development of the proof follows precisely the lines of the earlier case, § 1; and the two corollaries hold for the present case, Corollary 1 now taking the form:

$$\iint_S f(x, y)\, dS = f(\xi, \eta)\, A,$$

where (ξ, η) is a point of S, and A denotes the area of S.

Finally, consider the most general law of sub-division of S. Let n sub-regions be chosen as follows: (i) each sub-region shall conform to the general requirements imposed on the original region S, and no two sub-regions shall over-lap each other; moreover, no sub-region shall lie wholly outside of S; (ii) let S' be any region which, together with its boundary, lies wholly within S. Then for all values of n from a definite point on $(n \geqq N)$ the sub-regions shall cover all points of the region S'; (iii) the longest diameter of any sub-region corresponding to a given value of n shall approach 0 as its limit when n becomes infinite.

If, now, (x_k, y_k) be an arbitrary point of the k-th sub-region, and ΔS_k denote the area of the region, the sum

$$\sum_{k=1}^{n} f(x_k, y_k)\, \Delta S_k$$

will approach a limit as n becomes infinite, and the value of this limit is the double integral:

$$\lim_{n=\infty} \sum_{k=1}^{n} f(x_k, y_k)\, \Delta S_k = \iint_S f(x, y)\, dS.$$

Triple Integrals. The treatment here is precisely analogous. We begin with a finite region V of space, whose boundary is made up of a finite number of pieces of surfaces, each piece being capable of representation in at least one of the forms

$$z = \phi(x, y), \qquad y = \psi(x, z), \qquad x = \omega(y, z):$$

where the function on the right is continuous within and on the boundary of a suitably chosen region in the (x, y)-plane, or the (x, z)-plane, or the (y, z)-plane. Let $f(x, y, z)$ be a function continuous within and on the boundary of V.

We next divide space into equal cubes by the planes

$$x = \frac{i}{2^n}, \qquad y = \frac{j}{2^n}, \qquad z = \frac{k}{2^n},$$

where i, j, k, independently of each other, run through all integral values. Those cubes which lie wholly within V shall belong to Class (a); those which have at least one boundary point of V in their interior or on their boundary shall belong to Class (b).

Consider a cube of Class (a) or Class (b), and let (x_k, y_k, z_k) be a point of V which lies within or on the boundary of this cube. Form the sum

$$S_n = \sum f(x_k, y_k, z_k)\, \Delta V_k,$$

where ΔV_k denotes the volume of a cube (and is the same for all cubes, namely, 2^{-3n}), and the summation is extended to all cubes of Class (a), and to some, all, or none of the cubes of Class (b). *Then this sum approaches a limit as n becomes infinite.* The value of the limit is defined as the value of the *volume integral* of $f(x, y, z)$, extended throughout V:

$$\lim_{n=\infty} S_n = \iiint\limits_V f(x, y, z)\, dV.$$

The proof is given, as in the earlier case, by means of the sums

$$\sum{}^{+} M_k\, \Delta V_k, \qquad \sum{}^{-} m_k\, \Delta V_k.$$

Next, the two corollaries are deduced; and, finally, it is shown that, for the most general law of sub-division, the corresponding sum approaches a limit, and that the value of this limit is the triple integral:

$$\lim_{n=\infty} \sum f(x_k, y_k, z_k)\, \Delta V_k = \iiint\limits_V f(x, y, z)\, dV.$$

3. Iterated Integrals and the Fundamental Theorem. Let a region S of the (x, y)-plane be bounded by two curves:

(1) $$y = \Omega(x), \qquad y = \omega(x),$$

where $\Omega(x)$, $\omega(x)$, are both continuous in the interval $a \leqq x \leqq b$, and

$$\omega(x) < \Omega(x) \quad \text{if} \quad a < x < b.$$

In case $\omega(a) < \Omega(a)$, the segment of the line $x = a$, for which

$$\omega(a) < y < \Omega(a)$$

shall also belong to the boundary; and similarly if $\omega(b) < \Omega(b)$.

Let $f(x, y)$ be continuous within and on the boundary of S.

By the *iterated integral* of $f(x, y)$, extended over the region S:

$$(1) \qquad \int_a^b dx \int_{Y_0}^{Y_1} f(x, y)\, dy,$$

is meant the following. Let x be given any value in the interval $a \leqq x \leqq b$ and then held fast. Let

$$Y_0 = \omega(x), \qquad Y_1 = \Omega(x).$$

Form the integral

$$\int_{Y_0}^{Y_1} f(x, y)\, dy.$$

The value of this integral depends solely on the value of x which was chosen; *i.e.* it is a function of x. That the function is continuous is shown in Chap. XIX, § 1.

This function is now integrated over the interval (a, b), and the result is the iterated integral we set out to define.

Evaluation of the Double Integral. Let the plane be divided into squares as in § 2, the axis of x being divided at the same time into segments as in § 1; and let x_k be one of the points of division of the x-axis, which lies within the interval (a, b).

Consider the squares of Class (a) which have the line $x = x_k$ for their left-hand boundary. Add to these two rectangles as shown in the figure; *i.e.* the left-hand boundary of the rectangle shall have an extremity on the boundary of S. In particular, there may be no squares for a given k, and then we have just one rectangle, with the two ends of its left-hand boundary on the boundary of S. These squares and rectangles shall be taken as the sub-regions for the double integral

$$\iint_S f(x, y)\, dS,$$

the point at which $f(x, y)$ is to be formed being defined as follows.

Consider the square or rectangle whose horizontal sides lie in the lines

$$y = y_l, \qquad y = y_{l+1}.$$

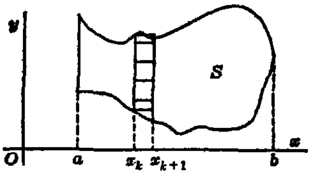


FIG. 64

Form the integral

$$\int_{y_l}^{y_{l+1}} f(x_k, y)\, dy.$$

By the Law of the Mean its value is

$$(y_{l+1} - y_l) f(x_k, y_l'); \qquad y_l < y_l' < y_{l+1}.$$

The point of this square or rectangle at which $f(x, y)$ is to be formed shall be (x_k, y_l').

The sum whose limit is the double integral can then be written in the form

$$(2) \qquad \sum f(x_k, y_l')\, \Delta S_{kl},$$

where, for each k, l is to run through the values corresponding to the squares and rectangles which abut from the right on the line $x = x_k$; and then these sums are to be added for $k = 1, 2$, etc.

On the other hand, consider the value of one of the above sums for a given k. Since

$$\Delta S_{kl} = \Delta x_k (y_{l+1} - y_l),$$

we have

$$\sum_l f(x_k, y_l')\, \Delta S_{kl} = \Delta x_k \sum_l \int_{y_l}^{y_{l+1}} f(x_k, y)\, dy = \Delta x_k \int_{Y_0}^{Y_1} f(x_k, y)\, dy$$

where $\qquad Y_0 = \omega(x_k), \qquad Y_1 = \Omega(x_k).$

Hence the total sum for all the squares and rectangles is

$$(3) \qquad \sum_k \Delta x_k \int_{Y_0}^{Y_1} f(x_k, y)\, dy.$$

The limit of this sum is the iterated integral (1) above.

We are now in a position to prove the equality of the double integral and the iterated integral:

$$\iint_S f(x, y)\, dS = \int_a^b dx \int_{Y_0}^{Y_1} f(x, y)\, dy.$$

The sums (2) and (3) are equal for all values of n. As n becomes infinite, the sum (2) approaches the double integral, the sum (3) the iterated integral, and the proof is complete.

The boundary of such a region S as that of § 2 may be cut by a parallel to the axis of y in four or more points. It is sufficient, however, for the needs of practice to restrict ourselves to such regions S as can be cut up into a finite number of regions, $S_1, S_2, \cdots$, each of which is of the type assumed in this paragraph. For any such region S, the double integral can be expressed as the sum of the iterated integrals taken over the regions $S_1, S_2, \cdots$.

The result here established is the *Fundamental Theorem of the Integral Calculus*, as stated in Chap. III, § 4. The first proof given in that earlier chapter, § 3, was based on geometric intuition. The second proof, § 17, was arithmetic, and it set forth the leading ideas of the argument, but it did not profess to carry through all the details. The present proof supplements the second one and leaves nothing to be desired in point of rigor. If the new definition of the definite integral, as given in §§ 1, 2, is once adopted, this proof is even simpler than the former proof.

Triple Integrals. The treatment admits of immediate extension to triple integrals, and thus we have a proof of the *Fundamental Theorem* in this case, namely, that (Chap. IV, § 2)

$$\iiint\limits_V f(x, y, z)\, dV = \iint\limits_S dS \int_{z_0}^{z_1} f(x, y, z)\, dz,$$

where S is the region of the (x, y)-plane whose points are the projections of the points of V, and the region V is bounded by the surfaces *

$$Z_0 = \omega(x, y), \qquad Z_1 = \Omega(x, y).$$

EXERCISE

Extend the treatment to a quadruple integral:

$$\iiiint\limits_R f(x, y, z, t)\, dR,$$

stating arithmetically the meaning of the geometric analogies.

* There is here a certain further restriction on V, which is not embarrassing in practice, since the projection S of V on the (x, y)-plane must now conform to the restrictions imposed on S in § 2.

4. Transformation of the Double Integral. Let the region S of § 2 be transformed on a region $\mathfrak{S}$ of the (u, v)-plane by means of the equations:

$$(1) \qquad \begin{cases} u = g(x, y) \\ v = h(x, y) \end{cases} \qquad \text{or} \qquad \begin{cases} x = G(u, v) \\ y = H(u, v) \end{cases}$$

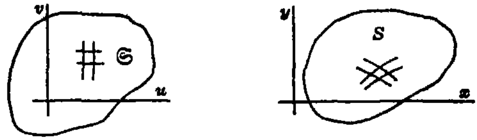


FIG. 65

the transformation being required to be one-to-one and continuous. Moreover, besides the existence and continuity of the first derivatives of g and h we require that the Jacobian

$$j = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix} = \frac{\partial(u, v)}{\partial(x, y)}$$

be different from zero at all points of S. The functions G, H will then also have continuous derivatives of the first order throughout $\mathfrak{S}$, and their Jacobian will likewise be different from zero at every point of $\mathfrak{S}$. For, $J = 1/j$, where

$$J = \frac{\partial(x, y)}{\partial(u, v)}.$$

Consider the double integral

$$(2) \qquad \iint_S f(x, y)\, dS.$$

Into what does this go when the transformation (1) is performed? Surely *not*, in general, into

$$\iint_{\mathfrak{S}} F(u, v)\, d\mathfrak{S},$$

where $$f(x, y) = F(u, v).$$

For, the latter integral is

$$\lim_{n=\infty} \sum F(u_k, v_k)\, \Delta\mathfrak{S}_k.$$

Now, a small region in $\mathfrak{S}$, — for example, a minute square, — goes over by (1) into a small region of S, whose shape is in general distorted, and whose area is not an equivalent infinitesimal, — as is shown by the simple example:

$$\begin{cases} u = 2x \\ v = y \end{cases} \qquad \begin{cases} x = \frac{1}{2}u \\ y = v \end{cases}$$

Here, $\Delta\mathfrak{S} = 2\,\Delta S$.

We can get some light, however, on what to expect by taking $\Delta\mathfrak{S}$ as a small square, bounded by the lines

$$u = u_0, \qquad u = u_0 + \Delta u, \qquad v = v_0, \qquad v = v_0 + \Delta v.$$

These lines go over in general into curved lines in the (x, y)-plane, and thus we have a curvilinear quadrilateral which, when the square is small, will look much like a parallelogram.

We can approximate to the area of the latter as follows. Recall the formula for the area of a triangle whose vertices are at (x_0, y_0), (x_1, y_1), (x_2, y_2). Aside from sign, it is

$$\frac{1}{2}\begin{vmatrix} x_0 & y_0 & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{vmatrix}.$$

The area of a parallelogram with three of its vertices in these three points will be twice as great.

Let (x_0, y_0) correspond to (u_0, v_0);
$(x_1, y_1) = (x_0 + \Delta_u x,\ y_0 + \Delta_u y)$ correspond to $(u_0 + \Delta u,\ v_0)$;
$(x_2, y_2) = (x_0 + \Delta_v x,\ y_0 + \Delta_v y)$ correspond to $(u_0,\ v_0 + \Delta v)$.

We have, then, as the area δS of the parallelogram, three of whose vertices lie at three of the corners of the curvilinear quadrilateral:

$$\delta S = \begin{vmatrix} x_0 & y_0 & 1 \\ x_0 + \Delta_u x & y_0 + \Delta_u y & 1 \\ x_0 + \Delta_v x & y_0 + \Delta_v y & 1 \end{vmatrix} = \begin{vmatrix} x_0 & y_0 & 1 \\ \Delta_u x & \Delta_u y & 0 \\ \Delta_v x & \Delta_v y & 0 \end{vmatrix} = \begin{vmatrix} \dfrac{\Delta_u x}{\Delta u} & \dfrac{\Delta_u y}{\Delta u} \\ \dfrac{\Delta_v x}{\Delta v} & \dfrac{\Delta_v y}{\Delta v} \end{vmatrix} \Delta u\, \Delta v$$

Since $\Delta\mathfrak{S} = \Delta u\,\Delta v$, it is clear that

$$\lim \frac{\delta S}{\Delta\mathfrak{S}} = \lim_{(\Delta u, \Delta v) \doteq (0, 0)} \begin{vmatrix} \dfrac{\Delta_u x}{\Delta u} & \dfrac{\Delta_u y}{\Delta u} \\ \dfrac{\Delta_v x}{\Delta v} & \dfrac{\Delta_v y}{\Delta v} \end{vmatrix} = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial y}{\partial u} \\ \dfrac{\partial x}{\partial v} & \dfrac{\partial y}{\partial v} \end{vmatrix} = \frac{\partial(x, y)}{\partial(u, v)} = J.$$

Now, it seems highly plausible geometrically that the area δS of the parallelogram and the area ΔS of the curvilinear quadrilateral *are* equivalent infinitesimals *; *i.e.* that

$$\lim \frac{\delta S}{\Delta S} = 1.$$

Hence follows:

$$\lim \frac{\Delta S}{\Delta \mathfrak{S}} = J = \frac{\partial(x, y)}{\partial(u, v)},$$

or

$$\Delta S = J \Delta \mathfrak{S} + \zeta \Delta \mathfrak{S},$$

where ζ is an infinitesimal.

We are now in a position to answer the question proposed at the start: What does the double integral

$$\int_S\int f(x, y)\, dS$$

become when referred to the transformed region $\mathfrak{S}$? Divide the (u, v)-plane into small squares, as in § 2. To the k-th square in $\mathfrak{S}$, whose area shall be denoted by $\Delta\mathfrak{S}_k$, corresponds a curvilinear quadrilateral in S, whose area shall be denoted by ΔS_k. If, now, (x_k, y_k) be any point of the latter region, and (u_k, v_k) the corresponding point of the former, then

$$\Delta S_k = J_k \Delta \mathfrak{S}_k + \zeta_k \Delta \mathfrak{S}_k,$$

where J_k denotes the value of J at (u_k, v_k). Hence

$$(3) \qquad \sum f(x_k, y_k)\, \Delta S_k = \sum F(u_k, v_k)[J_k \Delta \mathfrak{S}_k + \zeta_k \Delta \mathfrak{S}_k].$$

Let n become infinite. The limit of the left-hand side of (3) is the double integral (2). To the limit of the sum on the right we may apply Duhamel's Theorem, setting

$$\alpha_k = F(u_k, v_k)\, J_k \Delta \mathfrak{S}_k, \qquad \beta_k = F(u_k, v_k)[J_k \Delta \mathfrak{S}_k + \zeta_k \Delta \mathfrak{S}_k];$$

then

$$\lim_{n=\infty} \frac{\beta_k}{\alpha_k} = \lim_{n=\infty} \frac{J_k + \zeta_k}{J_k} = 1.$$

* This is, indeed, a fact; but a direct proof, based on infinitesimals, cannot easily be given, and so we agree (*a*) to accept the geometric evidence in all its suggestiveness as making reasonable the true result; (*b*) to make our proof depend on other and simpler analytic methods, cf. § 5. — Throughout this paragraph J, J_k, and the expressions for δS in terms of determinants should be replaced by their numerical values.

Hence the limit of this sum is the same as

$$\lim \sum F(u_k, v_k) J_k \Delta \mathfrak{S}_k = \int\int_{\mathfrak{S}} F(u, v)\, J\, d\mathfrak{S}.$$

We have, then, the final result, that

$$(4) \qquad \int\int_S f(x, y)\, dS = \int\int_{\mathfrak{S}} F(u, v) \,|\, J(u, v) \,|\, d\mathfrak{S}.$$

This result is true; but the proof is incomplete, in that we have not shown that the hypotheses of Duhamel's Theorem are fulfilled. We have adduced geometrical evidence which makes highly plausible the correctness of these hypotheses; but that is not mathematical proof. In the next paragraph we will give a proof.

5. Continuation. Proof by Line Integrals. It is possible to obtain an expression for the area of the region S by means of the theorems of Chap. XI, § 4:

$$(1) \qquad \int\int_S \frac{\partial P}{\partial y}\, dS = -\int_C P\, dx;$$

$$(2) \qquad \int\int_S \left(\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial x}\right) dS = -\int_C P\, dx + Q\, dy.$$

If in (1) we set $P = y$, the equation yields as the area, A, of S:

$$(3) \qquad A = -\int_C y\, dx.$$

If in (2) we set $P = y, \ Q = -x$, we have

$$(4) \qquad A = -\frac{1}{2}\int_C y\, dx - x\, dy.$$

Let the boundary, C, be represented in parametric form by the equations:

$$(5) \qquad x = \Phi(\lambda), \qquad y = \Psi(\lambda), \qquad 0 \leq \lambda \leq 1.$$

Then (3) gives

$$(6) \qquad A = -\int_0^1 y \frac{dx}{d\lambda}\, d\lambda.$$

Consider what this integral becomes when we make the transformation (1) of § 4. First, the boundary C of S goes over into the boundary $\mathfrak{C}$ of $\mathfrak{S}$, whose analytic representation is obtained by substituting the values of x and y from (5) in the first pair of equations (1) of § 4. Let the result be written in the form

$$(7) \qquad u = \phi(\lambda), \qquad v = \psi(\lambda),$$

where $\quad \phi(\lambda) = g\{\Phi(\lambda), \Psi(\lambda)\}, \qquad \psi(\lambda) = h\{\Phi(\lambda), \Psi(\lambda)\}.$

Thus in equation (6) we have:

$$y = H(u, v), \qquad \frac{dx}{d\lambda} = \frac{\partial G}{\partial u}\phi'(\lambda) + \frac{\partial G}{\partial v}\psi'(\lambda),$$

$$-\int_0^1 y\frac{dx}{d\lambda}d\lambda = -\int_0^1 H\left[\frac{\partial G}{\partial u}\phi'(\lambda) + \frac{\partial G}{\partial v}\psi'(\lambda)\right]d\lambda,$$

or

$$(7) \qquad A = \mp \int_{\mathfrak{C}} H\frac{\partial G}{\partial u}du + H\frac{\partial G}{\partial v}dv,$$

where the upper sign holds when $\mathfrak{C}$ is described in the positive sense; otherwise, the lower.

This last integral can be transformed as follows. Write equation (2) for the region $\mathfrak{S}$:

$$(8) \qquad \iint_{\mathfrak{S}} \left(\frac{\partial \mathfrak{P}}{\partial v} - \frac{\partial \mathfrak{Q}}{\partial u}\right) d\mathfrak{S} = -\int_{\mathfrak{C}} \mathfrak{P}\,du + \mathfrak{Q}\,dv,$$

and set $\qquad \mathfrak{P} = H\dfrac{\partial G}{\partial u}, \qquad \mathfrak{Q} = H\dfrac{\partial G}{\partial v}.$

Then $\qquad \dfrac{\partial \mathfrak{P}}{\partial v} - \dfrac{\partial \mathfrak{Q}}{\partial u} = \dfrac{\partial G}{\partial u}\dfrac{\partial H}{\partial v} - \dfrac{\partial G}{\partial v}\dfrac{\partial H}{\partial u} = \dfrac{\partial(G, H)}{\partial(u, v)} = J(u, v).$

Hence,

$$-\int_{\mathfrak{C}} H\frac{\partial G}{\partial u}du + H\frac{\partial G}{\partial v}dv = \iint_{\mathfrak{S}} \frac{\partial(x, y)}{\partial(u, v)}d\mathfrak{S},$$

or

$$(9) \qquad A = \pm \iint_{\mathfrak{S}} J(u, v)\,d\mathfrak{S}.$$

Since A is necessarily positive, we see that a transformation (1), § 4 has a positive Jacobian when a positive description of C leads to a positive description of $\mathfrak{C}$; otherwise, a negative Jacobian.

Stated as a theorem the result is as follows:

THEOREM. *The area A of the region S is represented by the double integral of the numerical value of the Jacobian $J(u, v)$, extended over the region $\mathfrak{S}$,*

$$A = \iint\limits_{\mathfrak{S}} \left| \frac{\partial(x, y)}{\partial(u, v)} \right| d\mathfrak{S}.$$

It is now easy to obtain a rigorous deduction of equation (3), § 4. Since

$$\Delta S_k = \iint\limits_{\Delta\mathfrak{S}_k} |J(u, v)|\, d\mathfrak{S},$$

we need only apply the Law of the Mean to this integral, and we have:

$$\Delta S_k = |J(u_k, v_k)|\, \Delta\mathfrak{S}_k,$$

where (u_k, v_k) is a properly chosen point of the k-th sub-region of $\mathfrak{S}$. If, then, we form the function f at the corresponding point (x_k, y_k), we have:

$$f(x_k, y_k)\, \Delta S_k = F(u_k, v_k)\, |J(u_k, v_k)|\, \Delta\mathfrak{S}_k.$$

Hence $$\lim_{n=\infty} \sum f(x_k, y_k)\, \Delta S_k = \lim_{n=\infty} \sum F(u_k, v_k)\, |J(u_k, v_k)|\, \Delta\mathfrak{S}_k,$$

or $$\iint\limits_S f(x, y)\, dS = \iint\limits_{\mathfrak{S}} F(u, v)\, |J(u, v)|\, d\mathfrak{S}, \qquad \text{q. e. d.}$$

6. The Iterated Integral. The evaluation of the double integral

(1) $$\iint\limits_{\mathfrak{S}} F(u, v)\, J(u, v)\, d\mathfrak{S}$$

by means of the iterated integral is immediate. Its value is

(2) $$\int_\alpha^\beta du \int_{V_0}^{V_1} F(u, v)\, J(u, v)\, dv \quad \text{or} \quad \int_\gamma^\delta dv \int_{U_0}^{U_1} F(u, v)\, J(u, v)\, du.$$

Hence we have a new evaluation of the original double integral:

(3) $$\iint\limits_S f(x, y)\, dS = \int_\alpha^\beta du \int_{V_0}^{V_1} F(u, v) \left| \frac{\partial(x, y)}{\partial(u, v)} \right| dv.$$

This last formula admits a new and important interpretation. In § 4 we interpreted equations (1):

(4) $$\begin{cases} u = g(x, y) \\ v = h(x, y) \end{cases} \qquad \begin{cases} x = G(u, v) \\ y = H(u, v) \end{cases}$$

as a *transformation of the points* (x, y) *of* S *into the points* (u, v) *of a region* $\mathfrak{S}$ *of the* (u, v)*-plane.*

It is, however, possible to put a wholly different interpretation on these equations. *We may regard the point* (x, y) *as fixed and the equations* (4) *as assigning to it new coordinates,* (u, v). Thus if

$$x = r\cos\phi, \qquad y = r\sin\phi,$$

we should most naturally interpret these equations as transforming the coordinates of the point from Cartesian axes to polar coordinates.

From this point of view, then, we introduce a system of curvilinear coordinates by means of equations (4), the functions g and h satisfying all of the conditions imposed at the beginning of § 4; and we arrive at a new iterated integral and the evaluation of the given double integral contained in (3).

Example. Let us apply the result embodied in formula (3) to obtaining the iterated integral in polar coordinates. Here,

$$J = \begin{vmatrix} \dfrac{\partial x}{\partial r} & \dfrac{\partial y}{\partial r} \\ \dfrac{\partial x}{\partial \phi} & \dfrac{\partial y}{\partial \phi} \end{vmatrix} = \begin{vmatrix} \cos\phi & \sin\phi \\ -r\sin\phi & r\cos\phi \end{vmatrix} = r,$$

and

$$\iint_S f(x, y)\,dS = \int_\alpha^\beta d\phi \int_{R_0}^{R_1} F(r, \phi)\, r\, dr,$$

the familiar formula. We observe that the factor r which presented itself in the earlier deduction is nothing more or less than the Jacobian.

EXERCISES

1. A system of curvilinear coordinates in the first quadrant is given by the two families of confocal parabolas:

$$y^2 = -2ux + u^2, \qquad y^2 = 2vx + v^2.$$

Compute the moment of inertia about the origin, of the region S bounded by two parabolas from each family.

2. Find the centre of gravity of S, Question 1.

3. The area of a curved surface is given by the integral, Chap. III, § 11:

$$A = \iint_S \sec\gamma\, dx\, dy.$$

If the surface is given parametrically, cf. § 8, Formulas (3), (4), (5) and Chap. VI, § 1, Formulas (9) and (14), then $\sec\gamma = \pm D/j$. Show that, on transforming from the Cartesian coordinates (x, y) to the curvilinear coordinates (λ, μ) we have:

$$A = \int_\sigma\int D\,d\lambda\,d\mu.$$

7. Extension to Triple Integrals. Let the region V of § 2 be transformed on a region $\mathfrak{B}$ of the (u, v, w)-space by means of the equations

$$(1)\qquad \begin{cases} u = g(x, y, z) \\ v = h(x, y, z) \\ w = l(x, y, z) \end{cases} \qquad\qquad \begin{cases} x = G(u, v, w) \\ y = H(u, v, w) \\ z = L(u, v, w) \end{cases}$$

where the same conditions of continuity are imposed as in § 4, and where, moreover, as there, the Jacobian shall be different from zero at every point of V:

$$j(x, y, z) = \frac{\partial(u, v, w)}{\partial(x, y, z)} \neq 0.$$

Then will

$$J(u, v, w) = \frac{\partial(x, y, z)}{\partial(u, v, w)}$$

also be different from zero, for $J = 1/j$.

Consider the volume integral

$$(2)\qquad \iint_V\int f(x, y, z)\,dV.$$

Into what does it go when the transformation (1) is performed? Our guess from analogy would be that

$$(3)\quad \iint_V\int f(x, y, z)\,dx\,dy\,dz = \iint_{\mathfrak{B}}\int F(u, v, w)\left|\frac{\partial(x, y, z)}{\partial(u, v, w)}\right| du\,dv\,dw,$$

and this is right.

A first proof by infinitesimals and Duhamel's Theorem can be given precisely as in the case of double integrals, § 4. A six-sided figure is cut from V by the surfaces

$$u = u_0, \qquad v = v_0, \qquad w = w_0,$$
$$u = u_0 + \Delta u, \qquad v = v_0 + \Delta v, \qquad w = w_0 + \Delta w,$$

and this figure looks very much like an oblique parallelepiped when Δu, Δv, Δw are all small. Its volume, ΔV, therefore, is seen to differ only by an infinitesimal of higher order from the volume δV of the corresponding parallelepiped, whose four vertices lie at the points:

$$(x_0, y_0, z_0), \qquad (x_0 + \Delta_u x, y_0 + \Delta_u y, z_0 + \Delta_u z),$$
$$(x_0 + \Delta_v x, y_0 + \Delta_v y, z_0 + \Delta_v z), \qquad (x_0 + \Delta_w x, y_0 + \Delta_w y, z_0 + \Delta_w z).$$

The value of δV is given by the formula:

$$\pm\, \delta V = \begin{vmatrix} \Delta_u x & \Delta_u y & \Delta_u z \\ \Delta_v x & \Delta_v y & \Delta_v z \\ \Delta_w x & \Delta_w y & \Delta_w z \end{vmatrix} = \begin{vmatrix} \dfrac{\Delta_u x}{\Delta u} & \dfrac{\Delta_u y}{\Delta u} & \dfrac{\Delta_u z}{\Delta u} \\ \dfrac{\Delta_v x}{\Delta v} & \dfrac{\Delta_v y}{\Delta v} & \dfrac{\Delta_v z}{\Delta v} \\ \dfrac{\Delta_w x}{\Delta w} & \dfrac{\Delta_w y}{\Delta w} & \dfrac{\Delta_w z}{\Delta w} \end{vmatrix} \Delta u\, \Delta v\, \Delta w.$$

The product $\Delta u \Delta v \Delta w$ represents precisely the volume $\Delta\mathfrak{V}$ of the corresponding region of the (u, v, w)-space, and so we have

$$\pm \lim \frac{\delta V}{\Delta\mathfrak{V}} = \frac{\partial(x, y, z)}{\partial(u, v, w)}.$$

Hence
$$\Delta V = |J|\, \Delta\mathfrak{V} + \zeta \Delta\mathfrak{V},$$

when ζ is infinitesimal, and so, by Duhamel's Theorem,

$$\lim \sum f(x_k, y_k, z_k)\, \Delta V_k = \lim \sum F(u_k, v_k, w_k)\, |J(u_k, v_k, w_k)|\, \Delta\mathfrak{V}_k,$$

or
$$\iiint\limits_V f(x, y, z)\, dV = \iiint\limits_{\mathfrak{V}} F(u, v, w)\, |J(u, v, w)|\, d\mathfrak{V}, \quad \text{q. e. d.}$$

The incompleteness in the proof lies in the fact that we cannot give a rigorous demonstration that the hypothesis of Duhamel's Theorem:

$$\beta_k = \alpha_k + \zeta_k \alpha_k, \qquad |\epsilon_k| < \eta,$$

is fulfilled. The geometric picture makes highly plausible the result, but we cannot be sure without proof that the geometric picture we see always represents the facts.

8. Continuation. Proof by Surface Integrals. The theorems of Chap. XI, § 9 yield the following expressions for the volume of the region V. On setting $C = z$ in (2), we have:

(1)
$$V = \iint\limits_S z \cos \gamma\, dS.$$

Secondly, in Equation I. of that same paragraph, let $A = x$, $B = y$, $C = z$; then

(2) $$V = \tfrac{1}{3}\int\!\!\int_S (x\cos\alpha + y\cos\beta + z\cos\gamma)\, dS.$$

Let the boundary of V, *i.e.* the surface S, be represented parametrically by the equations:

(3) $$x = \Phi(\lambda, \mu), \qquad y = \Psi(\lambda, \mu), \qquad z = \Omega(\lambda, \mu),$$

where the functions on the right are continuous, together with their first derivatives, and not all the Jacobians

(4) $$j_1 = \frac{\partial(y, z)}{\partial(\lambda, \mu)}, \qquad j_2 = \frac{\partial(z, x)}{\partial(\lambda, \mu)}, \qquad j_3 = \frac{\partial(x, y)}{\partial(\lambda, \mu)}$$

are zero at any one point; cf. Chap. VI, § 1. Let

(5) $$D = \sqrt{j_1^2 + j_2^2 + j_3^2}.$$

Then the area of a portion of S is given by the formula, cf. § 6, Ex. 3:

(6) $$S = \int\!\!\int_\sigma D\, d\lambda\, d\mu,$$

extended over the corresponding region σ of the (λ, μ)-plane. Moreover, the direction cosines of the outer normal of S are given by the formulas

(7) $$\cos\alpha = \frac{j_1}{D}, \qquad \cos\beta = \frac{j_2}{D}, \qquad \cos\gamma = \frac{j_3}{D},$$

provided the parameters λ, μ are suitably chosen.

Thus formula (1) can be written as follows:

(8) $$V = \int\!\!\int_\sigma z\,\frac{\partial(x, y)}{\partial(\lambda, \mu)}\, d\sigma,$$

where the integration is to be extended over the whole surface S.

We proceed to transform this integral to the (u, v, w)-space and the surface $\mathfrak{S}$ which is the image of S. A parametric representation is obtained for $\mathfrak{S}$ by substituting in the first set of equations (1), § 7, for x, y, z the values given by (3):

$$u = g\{\Phi(\lambda, \mu), \Psi(\lambda, \mu), \Omega(\lambda, \mu)\}, \qquad v = \text{etc.}$$

Let the result be written in the form

(9) $$u = \phi(\lambda, \mu), \qquad v = \psi(\lambda, \mu), \qquad w = \omega(\lambda, \mu).$$

The Jacobian which enters in (8) is seen to have the following value:

$$\begin{vmatrix} \frac{\partial x}{\partial \lambda} & \frac{\partial y}{\partial \lambda} \\ \frac{\partial x}{\partial \mu} & \frac{\partial y}{\partial \mu} \end{vmatrix} = \begin{vmatrix} G_1 \frac{\partial u}{\partial \lambda} + G_2 \frac{\partial v}{\partial \lambda} + G_3 \frac{\partial w}{\partial \lambda} & H_1 \frac{\partial u}{\partial \lambda} + H_2 \frac{\partial v}{\partial \lambda} + H_3 \frac{\partial w}{\partial \lambda} \\ G_1 \frac{\partial u}{\partial \mu} + G_2 \frac{\partial v}{\partial \mu} + G_3 \frac{\partial w}{\partial \mu} & H_1 \frac{\partial u}{\partial \mu} + H_2 \frac{\partial v}{\partial \mu} + H_3 \frac{\partial w}{\partial \mu} \end{vmatrix}$$

$$= \frac{\partial (G, H)}{\partial (v, w)} \frac{\partial (v, w)}{\partial (\lambda, \mu)} + \frac{\partial (G, H)}{\partial (w, u)} \frac{\partial (w, u)}{\partial (\lambda, \mu)} + \frac{\partial (G, H)}{\partial (u, v)} \frac{\partial (u, v)}{\partial (\lambda, \mu)}.$$

On substituting this value in (8) and recalling formulas (4), (5), (6), and (7), we have:

$$V = \pm \iint_{\mathfrak{S}} \left(L \frac{\partial (G, H)}{\partial (v, w)} \cos \alpha' + L \frac{\partial (G, H)}{\partial (w, u)} \cos \beta' + L \frac{\partial (G, H)}{\partial (u, v)} \cos \gamma' \right) d\mathfrak{S},$$

where α', β', γ' denote the angles made by the outer normal of $\mathfrak{S}$ with the axes of u, v, w respectively, and the $\pm$ sign must be so determined as to make the right-hand side positive.

This latter integral can be transformed into a volume integral by means of Equation I. of Chap. XI, § 9, written for the (u, v, w)-space:

$$\iiint_{\mathfrak{B}} \left(\frac{\partial \mathfrak{A}}{\partial u} + \frac{\partial \mathfrak{B}}{\partial v} + \frac{\partial \mathfrak{C}}{\partial w} \right) d\mathfrak{B} = \iint_{\mathfrak{S}} (\mathfrak{A} \cos \alpha' + \mathfrak{B} \cos \beta' + \mathfrak{C} \cos \gamma') \, d\mathfrak{S}.$$

Let $\mathfrak{A} = L \dfrac{\partial (G, H)}{\partial (v, w)}, \qquad \mathfrak{B} = L \dfrac{\partial (G, H)}{\partial (w, u)}, \qquad \mathfrak{C} = L \dfrac{\partial (G, H)}{\partial (u, v)}.$

Then
$$\frac{\partial \mathfrak{A}}{\partial u} + \frac{\partial \mathfrak{B}}{\partial v} + \frac{\partial \mathfrak{C}}{\partial w} = \frac{\partial (G, H, L)}{\partial (u, v, w)},$$

and hence
$$V = \iiint_{\mathfrak{B}} \left| \frac{\partial (x, y, z)}{\partial (u, v, w)} \right| d\mathfrak{B}.$$

From this point on the analysis is precisely like that of § 5 for double integrals and hence formula (3) of § 7 is established, or:

$$\iiint_{V} f(x, y, z) \, dV = \iiint_{\mathfrak{B}} F(u, v, w) \left| \frac{\partial (x, y, z)}{\partial (u, v, w)} \right| d\mathfrak{B}.$$

The Iterated Integral. This last integral can be evaluated in the usual way by means of the iterated integral, and hence we have a new evaluation for the original integral by means of an iterated integral:

$$\iiint\limits_V f(x, y, z)\,dV = \iint\limits_{\mathfrak{S}} d\mathfrak{S} \int\limits_{W_0}^{W_1} F(u, v, w)\,|\,J(u, v, w)\,|\,dw,$$

where $\mathfrak{S}$ here denotes the projection of the points of $\mathfrak{V}$ on the (u, v)-plane.

As in the case of double integrals, so here we can give an entirely new interpretation to equations (1) of § 7, considering them not as transforming the point (x, y, z) into a new point (u, v, w) of space; but rather as assigning to the point (x, y, z), which now remains fixed, new curvilinear coordinates, (u, v, w).

Example. Let the Cartesian coordinates (x, y, z) be replaced by spherical coordinates; cf. *Analytic Geometry*, p. 584:

$$x = r \sin\phi \cos\theta, \qquad y = r \sin\phi \sin\theta, \qquad z = r\cos\phi,$$

where θ denotes the longitude and ϕ, the co-latitude. Then

$$J(r, \phi, \theta) = \frac{\partial(x, y, z)}{\partial(r, \phi, \theta)} = r^2 \sin\phi$$

and we have the usual formula:

$$\iiint\limits_V f(x, y, z)\,dV = \iint\limits_S dS \int\limits_{R_0}^{R_1} F(r, \phi, \theta)\, r^2 \sin\phi\, dr$$

$$= \int\limits_{\alpha}^{\beta} d\phi \int\limits_{\Theta_0}^{\Theta_1} d\theta \int\limits_{R_0}^{R_1} F(r, \phi, \theta)\, r^2 \sin\phi\, dr.$$

EXERCISE

1. Obtain the iterated integral in cylindrical coordinates

$$x = r\cos\phi, \qquad y = r\sin\phi, \qquad z = z:$$

$$\iiint\limits_V f(x, y, z)\,dV = \iint\limits_S dS \int\limits_{R_0}^{R_1} F(r, \phi, z)\, r\,dr,$$

and explain the double integral.

9. Application to Hydrodynamics and Elasticity. Consider any three-dimensional material substance, a region τ_0 of which is changed in size and shape, and thus comes to occupy a new region, τ. For example, think of the currents of air in the atmosphere. Let τ_0 be a sphere 100 ft. in diameter, situated half a mile above the earth's surface at the instant of time $t = t_0$; and think of what has become of this sphere 30 seconds later. Let τ denote its new form, quite different, in all probability, from a sphere. But we assume that no discontinuity has taken place, so that the points of τ are related in a one-to-one manner and continuously to the points of τ_0.

Let ρ denote the density at any point of τ, and ρ_0, the density at the corresponding point of τ_0. We will assume that both functions are continuous. In general, ρ will not be equal to ρ_0. We shall presently find, however, a relation between them.

The total mass of air enclosed in τ must be precisely the same as the total mass in τ_0. Hence we have the equation

(1) $$\iiint_{\tau_0} \rho_0 \, d\tau_0 = \iiint_{\tau} \rho \, d\tau.$$

Let the Cartesian coordinates of any point of τ be (a, b, c), and the coordinates of the corresponding point of τ_0 be (a_0, b_0, c_0). Then (a, b, c) are connected with (a_0, b_0, c_0) by three equations such as (1), § 7.

$$\left\{\begin{aligned} a_0 &= g(a, b, c) \\ b_0 &= h(a, b, c) \\ c_0 &= l(a, b, c) \end{aligned}\right. \qquad \left\{\begin{aligned} a &= G(a_0, b_0, c_0) \\ b &= H(a_0, b_0, c_0) \\ c &= L(a_0, b_0, c_0) \end{aligned}\right.$$

Let $$J(a, b, c) = \frac{\partial(a, b, c)}{\partial(a_0, b_0, c_0)}.$$

The integral on the right of (1), when transformed to the region τ_0, takes on the value:

(2) $$\iiint_{\tau} \rho \, d\tau = \iiint_{\tau_0} \rho J \, d\tau_0.$$

From (1) and (2) we infer that

$$\iiint_{\tau_0} \rho J \, d\tau_0 = \iiint_{\tau_0} \rho_0 \, d\tau_0,$$

or

(3) $$\iiint_{\tau_0} (\rho J - \rho_0) \, d\tau_0 = 0.$$

The region τ_0 was wholly arbitrary. We can understand, then, by τ_0 any sub-region of the original τ_0, and equation (3) will still hold for this new region. From this we can infer that *the integrand must vanish at every point of the original* τ_0:

$$\rho J - \rho_0 = 0.$$

For, suppose that this function of (a_0, b_0, c_0) were positive at some point, A, of the original τ_0. Then we could enclose A in a new region, τ_0', at every point of which $\rho J - \rho_0$ would be positive, since this function is continuous and so cannot abruptly change sign. But here is a contradiction, for the integral of a positive function is necessarily positive, and not zero. Similarly, if $\rho J - \rho_0$ were negative at any point of the original τ_0.

We have thus obtained a relation which holds between ρ and ρ_0 at every pair of corresponding points, namely:

$$(4) \qquad \rho J = \rho_0, \qquad \text{where} \qquad J = \frac{\partial(a, b, c)}{\partial(a_0, b_0, c_0)}.$$

This is one form of the *Equation of Continuity*, — one of the basal theorems of hydromechanics.

Elasticity. We have taken as our physical picture a gas, deformed in an easily imaginable manner. But the same analysis would obviously apply to a piece of rubber in the tire of an automobile, or a piece of steel in the drive shaft. Equation (4) holds equally in all these cases.

10. Flux across a Surface. Consider a fluid in motion. If we fix our attention on a region R within the region in which the flow is continuous, then at any given instant, $t = t_1$, each particle of the fluid is moving with a definite vector velocity, and the totality of these vectors constitutes what is known as a *vector field*.

We wish to find the rate at which the mass is passing across a given surface, Σ, (open or closed) lying in R. This is known as the *flux across* Σ.

A Suggestive Special Case. The simplest case is that in which the fluid is of constant density, ρ, and is frozen, and, moreover, is moving without rotation, and with pure translation. Here, the vector velocity of each point is the same as that of every other point, and does not change with the time. Let this vector be $\mathfrak{V}$; let its length be V $(=|\mathfrak{V}|)$; and let its components along the axes be u, v, w.

Furthermore, let Σ be a plane surface, of area σ, whose normal, $\mathfrak{N}$, has the direction angles α, β, γ.

The amount, ΔM, of fluid which crosses Σ in Δt seconds is here readily visualized. It is the mass of an oblique cylinder, whose base is Σ and whose altitude is the projection of $\mathfrak{V}\Delta t$ along the normal $\mathfrak{N}$:

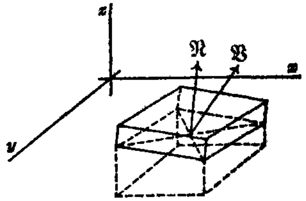


FIG. 66

(1) $$\Delta M = \rho(V\cos\epsilon)\,\sigma\,\Delta t.$$

Since

$$V\cos\epsilon = u\cos\alpha + v\cos\beta + w\cos\gamma,$$

we have

$$\frac{\Delta M}{\Delta t} = \rho(u\cos\alpha + v\cos\beta + w\cos\gamma)\,\sigma.$$

The flux is here constant, and $\lim \dfrac{\Delta M}{\Delta t} = \dfrac{\partial M}{\partial t}$. Hence

(2) $$\frac{\partial M}{\partial t} = \rho(u\cos\alpha + v\cos\beta + w\cos\gamma)\,\sigma.$$

The General Case. Returning now to the general case, consider a small piece of the region Σ, of area $\Delta\sigma$. This piece will look almost like a plane, of approximately the same area, $\Delta\sigma$; and, moreover, the vector velocities of the particles of the fluid near this piece will all be nearly equal to one another, so that the vector velocity $\mathfrak{V}$ of one particle, taken at random in Σ, will represent very closely the vector velocities of all its neighbors. Hence the rate of flow across $\Delta\sigma$ is given approximately by the formula

$$\rho(u\cos\alpha + v\cos\beta + w\cos\gamma)\,\Delta\sigma;$$

the normal to Σ being taken at the same point as the particle in question.

If we divide the whole region Σ up into n patches, assume an arbitrary point in each patch, denote the area of the patch by $\Delta\sigma_k$, and form the sum

(3) $$\Delta t \sum \rho_k(u_k\cos\alpha_k + v_k\cos\beta_k + w_k\cos\gamma_k)\,\Delta\sigma_k,$$

this sum will represent very closely the quantity ΔM of the fluid which has crossed Σ in the Δt seconds following the instant t (Δt also being assumed small). Hence the average rate of flow across Σ in the Δt seconds, or $\Delta M/\Delta t$, will be nearly equal to the sum

(4) $$\sum \rho_k(u_k\cos\alpha_k + v_k\cos\beta_k + w_k\cos\gamma_k)\,\Delta\sigma_k;$$

and the approximation will be better, the smaller the patches and the smaller Δt are taken. The limit of the above sum,

(5) $$\lim \sum \rho_k (u_k \cos \alpha_k + v_k \cos \beta_k + w_k \cos \gamma_k) \Delta \sigma_k,$$

when the largest diameter of any patch approaches 0 as its limit, and Δt also approaches 0, will be precisely the rate of flow, or *flux*, across Σ:

(6) $$\frac{\partial M}{\partial t} = \iint_{\Sigma} \rho (u \cos \alpha + v \cos \beta + w \cos \gamma)\, d\sigma.$$

Critique of the Foregoing. We have made a succession of positive statements, without attempting to give any other reason for their correctness than the physical picture. "Realize the physical situation and be convinced that these things are so," is the spirit of the text. Certainly, this is altogether properly the first step toward recognizing the reasonableness of the result, (6). But what are the physical laws we are assuming? Is the situation like that in which the area of a curved surface was formulated as an integral:

$$A = \iint_{S} \sec \gamma \, dS,$$

the final formula, — the integral itself, — being the simplest physical axiom to lay down? Or is it rather as it was in the case of fluid pressure, where physical laws of such simplicity that it seemed pedantic to state them led to the determination of the pressure:

$$P = w \int_{a}^{b} (x + c)\, y \, dx \qquad \text{or} \qquad P = w \iint_{S} (x + c)\, dS,$$

as a mathematical theorem?

The answer is as follows. The physical picture of the flow, which we have thrown on the screen of our imagination, though highly suggestive, is not sufficiently refined, in the absence of further elaboration, to render a mathematical deduction possible. It is precisely at this point that the infinitesimals of Leibniz, so dear to the heart of the mathematicians of the eighteenth century and to many a physicist of today, befuddle the situation, for they seem to go beyond the point we have reached above and to deliver a proof of (6), where our methods recognize their limitations.

It is, however, an error to attribute to them magical and mystical powers. They cover with a smoke screen the inherent difficulties; they do not surmount them.

A proof of (6) by infinitesimals is inconvenient, not for mathematical, but for physical reasons. It is not possible to frame a simple statement of physical facts which will yield the double inequality of Duhamel's Theorem, or its equivalent. The student has, then, two courses from which to choose. He may say: "The physical evidence is already highly convincing, though not of the order which we regard as ultimate. I prefer to accept the result and to go on and see what can be gotten out of it." This is the stand of a man who respects himself scientifically, but who does not wish, for the time being, at least, to study further the mathematical-physical situation.

The other course is to attempt a proof of (6) on the basis of clean-cut assumptions defining what we mean by a flow of a fluid. The proof is not a brief one; but the ideas and methods it involves, far from being artificial and developed for just this one case, are such as the student will meet again and again both in pure mathematics and in mathematical physics. For this reason a careful study of the proof will well reward the effort.

One other word. We have considered a material fluid. We might equally well think of the substance as heat or electricity. Even more generally, the results apply to the case of any vector field (as in the *flux of force* across a surface in electricity or magnetism or gravitation). There may be a point function, like the density ρ above, given at each point of the field; or such a function may be absent: $\rho = 1$.

11. Continuation. Proof of the Formula. *Definition of a Flow.* We must first define what we mean by the *flow* of a fluid. Let R_0 be a region contained in the fluid at time $t = t_0$. Consider an arbitrary particle which initially (*i.e.* when $t = t_0$) is at a point (a, b, c) of R_0. Let (x, y, z) be the point at which this particle has arrived at an arbitrary later instant, $t = t$. Then

$$(1)\qquad \begin{cases} x = f(a, b, c, t) \\ y = \phi(a, b, c, t) \\ z = \psi(a, b, c, t) \end{cases}$$

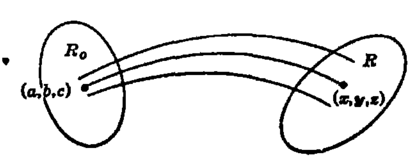


FIG. 67

where these functions are continuous. And now the region R_0 is carried over, in a one-to-one manner and continuously, into a region R, which changes with the time, but for a given value of t is altogether definite. In other words, equations (1) can be solved for a, b, c:

$$(2)\qquad \begin{cases} a = F(x, y, z, t) \\ b = \Phi(x, y, z, t) \\ c = \Psi(x, y, z, t) \end{cases}$$

where (x, y, z) is any point of the region R which corresponds to the value of t in question, and the functions F, Φ, Ψ are continuous.

Finally, we require that the functions f, ϕ, ψ have continuous first partial derivatives in all the arguments, and that the Jacobian be always positive:

$$(3)\qquad \frac{\partial(x, y, z)}{\partial(a, b, c)} > 0 \quad \text{throughout } R_0.$$

From these conditions it follows that the functions F, Φ, Ψ also have continuous first partial derivatives. Moreover,

$$(4)\qquad \frac{\partial(a, b, c)}{\partial(x, y, z)} > 0 \quad \text{throughout } R.$$

For, the Jacobian in (4) is the reciprocal of that in (3); cf. Chap. V, §§ 12, 13.

The Problem. We wish to determine the rate at which the fluid is traversing a given surface, open or closed, which lies in R. More precisely, if ΔM denotes the mass of the fluid which crosses this surface in the interval of time from $t = t_1$ to $t = t_1 + \Delta t$, then we wish to find the

$$(5)\qquad \lim_{\Delta t \doteq 0} \frac{\Delta M}{\Delta t}, \qquad \text{or} \qquad \frac{\partial M}{\partial t}.$$

The Surface Σ. Consider a piece, Σ, of this surface, and require that Σ can be represented in the form

$$(6)\qquad z = \omega(x, y),$$

where ω is continuous, together with its first derivatives, throughout the projection S of Σ on the (x, y)-plane.

Consider the instant $t = t_1$ and the succeeding interval

$$t_1 \leqq t \leqq t_1 + h.$$

At any instant, $t = \tau$, of this interval, there will be a certain surface of the particles coincident with Σ, and these will pass on as t increases, reaching a definite final position when $t = t_1 + h$. We will impose further conditions to make sure that no two members of this one-parameter family of surfaces ever have a point in common.

This can be done by requiring that the tangent to the path of any particle which at time $t = t_1$ is on Σ, be not tangent to Σ; and, secondly, by suitably restricting h.

Let us formulate this condition analytically. The components, u, v, w, of the velocity of any particle along the axes at any instant are the time derivatives of its coordinates, or

$$(7) \qquad u = f_4(a, b, c, t), \qquad v = \phi_4(a, b, c, t), \qquad w = \psi_4(a, b, c, t).$$

These can be taken as the direction components of the tangent to the path of the particle. The direction components of a normal to Σ are

$$\omega_1(x, y), \quad \omega_2(x, y), \quad -1.$$

For a particle at (x, y, z) on Σ, when $t = t_1$, we must then have

$$(8) \qquad u\omega_1 + v\omega_2 - w \neq 0,$$

where u, v, w, as given by (7), are formed for that point (a, b, c) of R_0 which corresponds to the above point (x, y, z) of Σ at time $t = t_1$.

The expression on the left of (8) can be regarded as a function of the independent variables (x, y) in S, and since it is continuous and different from zero, it will be either positive throughout S or else negative throughout S.

The Region $\mathfrak{R}$. We now have the physical picture of a three-dimensional distribution of the fluid which has passed across the surface Σ in the interval of time $(t_1, t_1 + h)$, and which, at $t = t_1 + h$, is spread out throughout a region $\mathfrak{R}$ of space, bounded in part by Σ. Is this physical picture justified by the hypotheses above laid down? This is purely a mathematical question, to the treatment of which we now turn.

It is convenient to represent the points of $\mathfrak{R}$ by means of the following system of curvilinear coordinates. Let the Cartesian coordinates of a point of S be denoted by (λ, μ). We thus have a simple system of curvilinear coordinates on the surface Σ, whose points (x, y, z) are represented by the equations

$$(9) \qquad x = \lambda, \qquad y = \mu, \qquad z = \omega(\lambda, \mu).$$

As the curvilinear coordinates (λ, μ, τ) of a particle in $\mathfrak{R}$, whose Cartesian coordinates are (x, y, z), we now choose the coordinates λ, μ of the point of Σ through which it passed, and the time $t = \tau$ of its passage. The expression of x, y, z as functions of λ, μ, τ, according to the above definition, is obtained as follows: (*i*) the point (a, b, c)

of R_0 corresponding to (x, y, z) of $\mathfrak{R}$ is given by the equations

$$(10)\quad \begin{cases} \lambda = f(a, b, c, \tau) \\ \mu = \phi(a, b, c, \tau) \\ \omega(\lambda, \mu) = \psi(a, b, c, \tau) \end{cases} \quad \text{or} \quad \begin{cases} a = F[\lambda, \mu, \omega(\lambda, \mu), \tau] \\ b = \Phi[\lambda, \mu, \omega(\lambda, \mu), \tau] \\ c = \Psi[\lambda, \mu, \omega(\lambda, \mu), \tau] \end{cases}$$

(*ii*) the values of a, b, c thus obtained are substituted in (1), and t is set $= t_1 + h$:

$$(11)\quad \begin{cases} x = f(a, b, c, t_1 + h) \\ y = \phi(a, b, c, t_1 + h) \\ z = \psi(a, b, c, t_1 + h) \end{cases}$$

We wish to show that the Jacobian

$$(12)\quad J = \frac{\partial(x, y, z)}{\partial(\lambda, \mu, \tau)} \neq 0$$

at every point of Σ when $\tau = t_1$ and $h = 0$; for it then follows from the implicit function theorem, Chap. V, § 12, that, when h is suitably restricted, equations (11) define a region $\mathfrak{R}$ such that, if (x, y, z) is an arbitrary point of $\mathfrak{R}$, these equations admit a unique solution, (λ, μ, τ), where λ, μ, τ are continuous functions of (x, y, z) in $\mathfrak{R}$, and moreover J will not vanish in $\mathfrak{R}$.

To compute J, combine equations (10) and (11), thus obtaining the following:

$$(13)\quad \begin{cases} x - \lambda = f(a, b, c, t_1 + h) - f(a, b, c, \tau) \\ y - \mu = \phi(a, b, c, t_1 + h) - \phi(a, b, c, \tau) \\ z - \omega(\lambda, \mu) = \psi(a, b, c, t_1 + h) - \psi(a, b, c, \tau) \end{cases}$$

It is now easy to compute the nine partial derivatives which enter in the Jacobian (12) *for the particular values* $\tau = t_1$ and $h = 0$. In computing the six of these derivatives that are taken with respect to λ and μ, it is allowable to set $\tau = t_1$ and $h = 0$ *before* the differentiation. The values of the remaining three, with respect to τ, are also computed with ease, and we have, as the final result:

$$\left.\frac{\partial(x, y, z)}{\partial(\lambda, \mu, \tau)}\right|_{\substack{\tau = t_1 \\ h = 0}} = \begin{vmatrix} 1 & 0 & -f_4(a, b, c, t_1) \\ 0 & 1 & -\phi_4(a, b, c, t_1) \\ \omega_1(\lambda, \mu) & \omega_2(\lambda, \mu) & -\psi_4(a, b, c, t_1) \end{vmatrix}$$

$$= u(a, b, c, t_1)\,\omega_1(\lambda, \mu) + v(a, b, c, t_1)\,\omega_2(\lambda, \mu) - w(a, b, c, t_1).$$

But this is precisely the left-hand side of (8), and hence is nowhere zero on Σ.

The Determination of ΔM. We are now in a position to determine the mass of the fluid contained in $\mathfrak{R}$. It is :*

$$M = \iiint\limits_{\mathfrak{R}} \rho \, dx \, dy \, dz = \iiint\limits_{\mathfrak{R}'} \rho \frac{\partial(x, y, z)}{\partial(\lambda, \mu, \tau)} \, d\lambda \, d\mu \, d\tau, \tag{14}$$

where $\mathfrak{R}'$ denotes the region of the space of the Cartesian coordinates (λ, μ, τ) which corresponds to $\mathfrak{R}$ by the transformation (11).

The volume integral can be evaluated by means of the iterated integral, and hence

$$M = \iint\limits_S dS \int\limits_{t_1}^{t_1+h} \rho \frac{\partial(x, y, z)}{\partial(\lambda, \mu, \tau)} \, d\tau. \tag{15}$$

Applying the law of the mean, we have:

$$\int\limits_{t_1}^{t_1+h} \rho \frac{\partial(x, y, z)}{\partial(\lambda, \mu, \tau)} \, d\tau = h \left[\rho \frac{\partial(x, y, z)}{\partial(\lambda, \mu, \tau)} \right]_{\tau = t_1 + \theta h}.$$

Since $J(\lambda, \mu, \tau)$ is continuous, the last factor differs uniformly from

$$\rho \frac{\partial(x, y, z)}{\partial(\lambda, \mu, \tau)} \Bigg|_{\tau = t_1}$$

by a quantity ζ which is infinitesimal with h. More precisely, the largest value that $|\zeta|$ has for any point (λ, μ) in S and for a given h approaches zero with h.

If, then, finally, we write M as ΔM and h as Δt, we have

$$\frac{\Delta M}{\Delta t} = \iint\limits_S \left[\rho \frac{\partial(x, y, z)}{\partial(\lambda, \mu, \tau)} \Bigg|_{\tau = t_1} + \zeta \right] dS. \tag{16}$$

Hence

$$\lim_{\Delta t \doteq 0} \frac{\Delta M}{\Delta t} = \frac{\partial M}{\partial t} = \iint\limits_S \rho \frac{\partial(x, y, z)}{\partial(\lambda, \mu, t_1)} \, dS. \tag{17}$$

But the value of the Jacobian J on Σ has been shown to be

$$u \omega_1 + v \omega_2 - w.$$

Consequently

$$\frac{\partial M}{\partial t} = \iint\limits_S \rho (u \omega_1 + v \omega_2 - w) \, dS. \tag{18}$$

* Save possibly as to sign. For a discussion of this question cf. below.

This latter integral can be written as a surface integral taken over Σ, and thus we have the familiar result:

$$\text{(19)} \qquad \frac{\partial M}{\partial t} = \iint\limits_{\Sigma} \rho(u \cos\alpha + v \cos\beta + w \cos\gamma)\, d\Sigma,$$

where α, β, γ are the direction angles of a normal to the surface.

The Signs. In formula (14) there should be a minus sign before the last integral if $J < 0$. Moreover, we have not said in (19) which normal is to be chosen. Nevertheless, (19) is true in all cases, no matter which normal be chosen, provided we take ΔM as an algebraic quantity, considering ΔM as positive when the direction of flow at any point of Σ makes an acute angle with the normal at that point, and negative in the other case.

The Excepted Case. It remains to consider the case that

$$\text{(20)} \qquad \Delta = u\omega_1 + v\omega_2 - w = 0$$

on Σ. This may happen either through the velocity vector being tangent to Σ or through the velocity vanishing on Σ.

Suppose the points of the first kind lie on a curve C and those of the second kind, on a curve C':

$$C: \qquad \Delta = 0; \quad (u, v, w) \neq (0, 0, 0);$$

$$C': \qquad \Delta = 0; \quad u = 0,\ v = 0,\ w = 0.$$

It might at first sight appear as if we could consider a portion Σ' of Σ not reaching up to either C or C' and, applying formula (19) to it, allow Σ' to approach Σ as its limit. Then the surface integral (19), extended over Σ', would approach the surface integral extended over Σ as its limit, and we should have the result we wish to establish. But there is a fallacy in this reasoning, which consists in inverting the order in a double limit.

Let $\Delta M'$ be the quantity of the fluid which flows across Σ' in Δt seconds and let $\Delta M''$ be the quantity which flows across the remainder of Σ in that time. Then the limit we wish to determine is

$$\frac{\partial M}{\partial t} = \lim_{\Delta t \to 0}\left(\frac{\Delta M'}{\Delta t} + \frac{\Delta M''}{\Delta t}\right) = \frac{\partial M'}{\partial t} + \frac{\partial M''}{\partial t}.$$

If now we let Σ' approach Σ as above, it is true that $\partial M'/\partial t$ approaches the limit given by formula (19). But

$$\frac{\partial M}{\partial t} = \lim \frac{\partial M'}{\partial t} + \lim \frac{\partial M''}{\partial t},$$

and why should this last limit be 0? In other words, why should

$$\lim_{\Sigma'\to\Sigma}\frac{\partial M''}{\partial t}=\frac{\partial}{\partial t}\left(\lim_{\Sigma'\to\Sigma}M''\right)?$$

It is this point which requires proof, and the proof can be given by direct appraisal of $\Delta M''$.

Let the curve C' (which may consist of several pieces) be embedded in one or more slender strips, Σ_1. Let C also be embedded in one or more slender strips, and denote the part of these strips which lies outside Σ_1 by Σ_2. Denote the quantity of fluid which flows across Σ_1 and Σ_2 in Δt seconds by ΔM_1 and ΔM_2.

To obtain an appraisal for ΔM_1, let V denote the largest value of the velocity which any particle, traversing a point P of Σ_1 at time $t=\tau$ intermediate between t_1 and $t_1+\Delta t$, attains in the interval from τ to $t_1+\Delta t$. Replace Δt for convenience by the former notation, h. Then the particles which traverse the point P in the interval of time from t_1 to t_1+h will lie in a sphere about P, of radius Vh.

Let P sweep out Σ_1, carrying such a sphere with it, and let the region of space swept out by the sphere be denoted by U_1. Then U_1 is of the nature of a shell, encasing Σ_1 and of thickness about $2Vh$. Its volume will be approximately $2VhA_1$, where A_1 is the area of Σ_1. It remains to show, with all rigor, that this volume is less than $2VhB_1$, where B_1 is a suitably chosen constant. This proof can be carried through without difficulty, and is left to the student. Thus

$$|\Delta M_1|<2VhB_1.$$

An appraisal of ΔM_2 can be obtained as follows. Consider a particle of the fluid which passes a point P of Σ_2 at any instant, $t=\tau$, of the interval from t_1 to t_1+h. Let δ' be the greatest angle which its path ever makes with the tangent plane at P during the interval from τ to t_1+h, and let V' be its greatest velocity in this interval. Let δ and V be respectively the maximum values of δ' and V', as P sweeps out Σ_2. Then Σ_2 can be so chosen that δ will be arbitrarily small. It is clear that those particles which pass a given point P of Σ_2 in the interval from t_1 to t_1+h will lie between two cones whose axes are in the normal to Σ at P and whose semi-vertical angle is $\frac{1}{2}\pi-\delta$. No particle will depart from P by more than Vh, and each particle will remain in the region R between the above cones and within a co-axial cylinder of radius Vh. The part of Σ within this cylinder will also lie in R, if h is suitably restricted.

Now, the distance of the most remote point of R from Σ is obviously less than $2Vh\tan\delta$, except when P is near the edge of Σ_2; and this exception can be conveniently removed by extending Σ_2 slightly, to form a region Σ_2', of area A_2.

If, now, we construct a sphere of radius $2Vh\tan\delta$ about P as centre, and allow P to describe Σ_2', carrying its sphere with it, a region U_2 of space will thus be swept out, the volume of which is approximately

$$4A_2Vh\tan\delta,$$

and it follows from the evaluation of the volume of U_1 in the earlier case that the volume of U_2 is less than $4B_2Vh\tan\delta$, where B_2 denotes a suitably chosen constant. Thus

$$|\Delta M_2| < 4B_2Vh\tan\delta.$$

We see, then, that Σ_1, Σ_2 can be so chosen that $\Delta M_1/\Delta t$ and $\Delta M_2/\Delta t$ remain numerically as small as one likes for all values h that are positive and small. Hence $\partial M''/\partial t$ does approach 0, and the proof is complete.

It is possible that the locus $\Delta = 0$ consists, not of curves C and C', but of whole two-dimensional regions. The above proof applies, however, to these cases without modification.

The result, namely, equation (19), is independent of the choice of the coordinate axes. Hence it holds for any bilateral surface which can be cut up into pieces, each of which, when referred to properly chosen axes, can be represented in the form (6).

12. The Equation of Continuity. Consider an arbitrary subregion, V, lying in the substance, the flow of which is defined above. The quantity of matter, M, in V is given by the integral:

(1) $$M = \iiint\limits_V \rho\, dV,$$

where ρ is computed, at the instant t in question, for each point (x, y, z) of the region V.

The rate at which M is increasing at this instant will be given, then, by the equation:

(2) $$\frac{\partial M}{\partial t} = \iiint\limits_V \frac{\partial \rho}{\partial t}\, dV,$$

ρ being a function of (x, y, z, t).

On the other hand, the rate at which M is increasing is the rate at which matter is flowing into V from without across the surface, S, of V, and thus is given by § 11 as

$$\frac{\partial M}{\partial t} = \iint_S \rho (u \cos\alpha + v \cos\beta + w \cos\gamma)\, dS, \tag{3}$$

where α, β, γ are the direction angles of the inner normal.

By Green's Theorem, Chap. XI, § 9, I., the right-hand side of (3) has the value*:

$$-\iiint_V \left\{ \frac{\partial(\rho u)}{\partial x} + \frac{\partial(\rho v)}{\partial y} + \frac{\partial(\rho w)}{\partial z} \right\} dV. \tag{4}$$

From these two values of $\partial M/\partial t$ we infer that

$$\iiint_V \left\{ \frac{\partial \rho}{\partial t} + \frac{\partial(\rho u)}{\partial x} + \frac{\partial(\rho v)}{\partial y} + \frac{\partial(\rho w)}{\partial z} \right\} dV = 0. \tag{5}$$

Now, the integrand is continuous throughout the whole region of flow, and the region V is arbitrary. It follows, then, that this function must vanish identically; cf. § 9 and Chap. XI, § 14 for the reasoning here employed:

$$\frac{\partial \rho}{\partial t} + \frac{\partial(\rho u)}{\partial x} + \frac{\partial(\rho v)}{\partial y} + \frac{\partial(\rho w)}{\partial z} = 0. \tag{6}$$

The result is known as the *Equation of Continuity*. It is a necessary and sufficient condition that nowhere in the region of flow is matter either generated or destroyed; there are no (three-dimensional) sources or sinks.

EXERCISE

Prove the Equation of Continuity by differentiating equation (4) of § 9 with respect to the time.

* We now make the further assumption that the functions in (1), § 11 possess continuous second partial derivatives.

CHAPTER XIII

VECTOR ANALYSIS

1. Vectors and their Addition. By a *vector* is meant a directed line segment, situated anywhere in space. Vectors will usually be denoted by German letters or by parentheses; thus a vector angular velocity may be written (ω).

Two vectors, $\mathfrak{A}$ and $\mathfrak{B}$, are defined as *equal* if they are parallel and have the same sense, and moreover are of equal length. We write:

$$\mathfrak{A} = \mathfrak{B}.$$

By the *absolute value* of a vector $\mathfrak{A}$ is meant its length; it is denoted by $|\mathfrak{A}|$.

Addition. By the *sum* of two vectors, $\mathfrak{A}$ and $\mathfrak{B}$, is meant their geometric sum, or the vector $\mathfrak{C}$ obtained by the parallelogram law. We write:

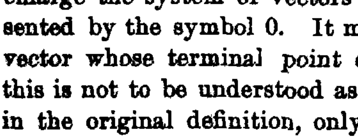


FIG. 68

$$\mathfrak{A} + \mathfrak{B} = \mathfrak{C}.$$

In order that this definition may apply in all cases, it is necessary to enlarge the system of vectors above defined by a *nul vector*, represented by the symbol 0. It may be thought of as a point, or as a vector whose terminal point coincides with its initial point; but this is not to be understood as meaning that it really was included in the original definition, only we were not shrewd enough to see it. It is a new element, added to the original system by a new and independent definition.

If $\mathfrak{B}$ is parallel to $\mathfrak{A}$ and of the same length, but opposite in sense, we write:

$$\mathfrak{A} + \mathfrak{B} = 0, \qquad \text{or} \qquad \mathfrak{B} = -\mathfrak{A}.$$

Moreover, we understand by $m\mathfrak{A}$, where m is any number* a

* By *number* we mean an ordinary real number, positive, negative or zero; rational or irrational.

vector parallel to $\mathfrak{A}$ and $|m|$-times as long; its sense being the same as that of $\mathfrak{A}$, or opposite, according as m is positive or negative. If $m = 0$, then $m\mathfrak{A}$ is a null vector:* $0\mathfrak{A} = 0$. The notation $\mathfrak{A}m$ means $m\mathfrak{A}$, and also

$$\frac{a\mathfrak{A} + b\mathfrak{B}}{a+b} \quad \text{means} \quad \frac{a}{a+b}\mathfrak{A} + \frac{b}{a+b}\mathfrak{B}.$$

Vector addition obeys the *commutative* and the *associative law* of ordinary algebra:

$$\mathfrak{A} + \mathfrak{B} = \mathfrak{B} + \mathfrak{A},$$

$$\mathfrak{A} + (\mathfrak{B} + \mathfrak{C}) = (\mathfrak{A} + \mathfrak{B}) + \mathfrak{C}.$$

Subtraction. By $\mathfrak{B} - \mathfrak{A}$ is meant that vector, $\mathfrak{X}$, which added to $\mathfrak{A}$ will give $\mathfrak{B}$:

$$\mathfrak{A} + \mathfrak{X} = \mathfrak{B}, \qquad \mathfrak{X} = \mathfrak{B} - \mathfrak{A}.$$

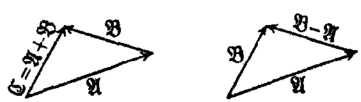


FIG. 69

It is easy to see how to obtain $\mathfrak{X}$ geometrically: Construct $\mathfrak{A}$ and $\mathfrak{B}$ with the same initial point; then $\mathfrak{B} - \mathfrak{A}$ is the vector whose initial point is the terminal point of $\mathfrak{A}$, and whose terminal point is the terminal point of $\mathfrak{B}$.

Cartesian Representation of a Vector. Let a system of Cartesian axes be chosen, and let $\mathfrak{i}$, $\mathfrak{j}$, $\mathfrak{k}$ be three unit vectors lying along these axes. Let $\mathfrak{A}$ be an arbitrary vector, whose components along the axes are X, Y, Z. Then evidently

$$\mathfrak{A} = X\mathfrak{i} + Y\mathfrak{j} + Z\mathfrak{k}.$$

If $$\mathfrak{B} = X'\mathfrak{i} + Y'\mathfrak{j} + Z'\mathfrak{k},$$

then $$\mathfrak{A} + \mathfrak{B} = (X + X')\mathfrak{i} + (Y + Y')\mathfrak{j} + (Z + Z')\mathfrak{k}.$$

Also: $$|\mathfrak{A}| = \sqrt{X^2 + Y^2 + Z^2}.$$

A point in space, with the coordinates (x, y, z), may be represented by the vector

$$\mathfrak{r} = x\mathfrak{i} + y\mathfrak{j} + z\mathfrak{k}.$$

If P_1 and P_2, represented by $\mathfrak{r}_1$ and $\mathfrak{r}_2$, are any two points in space, the mid-point of the line segment joining them is given by the terminal point of

* It is true that the symbol 0 is used in this equation in two different senses, — once, as the *number* 0, and again as a *nul vector*. This double use will not be found, however, to lead to confusion.

(1) $$\bar{\mathfrak{r}} = \frac{\mathfrak{r}_1 + \mathfrak{r}_2}{2},$$

where $\bar{\mathfrak{r}}$ is drawn from the origin as the initial point.

Let n masses, $m_1, \cdots, m_n$, be located anywhere in space, and let the coordinates of the k-th of these masses be (x_k, y_k, z_k). Then their centre of gravity is given by the terminal point of the vector

(2) $$\bar{\mathfrak{r}} = \frac{m_1\mathfrak{r}_1 + m_2\mathfrak{r}_2 + \cdots + m_n\mathfrak{r}_n}{m_1 + m_2 + \cdots + m_n},$$

the initial point being taken at the origin.

Equations (1) and (2) merely compress into a single vector equation what is expressed in ordinary form through three equations.

Vectors in Physics. Those physical quantities, like forces and velocities, which require for their expression, beside their magnitude, their direction and sense, can be represented by vectors. We may mention accelerations, couples, angular velocities and momentum, and the vector moment of a force about a point. The law of composition is in each case the law of vector addition given above. But this is not true of all physical quantities that can be represented by vectors. Thus any displacement of a sphere whose centre remains fixed is a vector quantity, but the displacement which arises as the result of two given displacements is not in general the displacement which corresponds to the sum of the vectors representing the given displacements.

EXERCISES

1. Show how to construct geometrically the sum of n vectors,

$$\mathfrak{A}_1 + \mathfrak{A}_2 + \cdots + \mathfrak{A}_n,$$

by means of a skew polygon, the generalization of Fig. 69.

2. If n forces, acting at a point, are represented by the n vectors $\mathfrak{F}_1, \mathfrak{F}_2, \cdots, \mathfrak{F}_n$, show that their resultant is represented by the vector

$$\mathfrak{F} = \mathfrak{F}_1 + \mathfrak{F}_2 + \cdots + \mathfrak{F}_n,$$

drawn from the point of application of the forces.

3. If a couple in space is represented by a vector, show that the resultant $\mathfrak{M}$ of n couples, $\mathfrak{M}_1, \cdots, \mathfrak{M}_n$, is represented by the vector *

$$\mathfrak{M} = \mathfrak{M}_1 + \mathfrak{M}_2 + \cdots + \mathfrak{M}_n.$$

* This question involves a knowledge only of the elementary theory of the composition and the resolution of couples in space.

4. Three vectors are said to be *complanar* * if there is a plane in space to which they are all parallel. Show that, if $\mathfrak{A}$, $\mathfrak{B}$, $\mathfrak{C}$ be three non-complanar vectors, any vector, $\mathfrak{X}$, can be written in the form

$$\mathfrak{X} = l\mathfrak{A} + m\mathfrak{B} + n\mathfrak{C},$$

and that l, m, n are uniquely determined.

5. If

$$\mathfrak{A} = \mathfrak{i}f(t) + \mathfrak{j}\phi(t) + \mathfrak{k}\psi(t),$$

show that a necessary and sufficient condition that $\mathfrak{A}$ have a derivative is that each of the functions $f(t)$, $\phi(t)$, $\psi(t)$ have a derivative. Then,

$$D_t\mathfrak{A} = \mathfrak{i}f'(t) + \mathfrak{j}\phi'(t) + \mathfrak{k}\psi'(t); \qquad d\mathfrak{A} = D_t\mathfrak{A}\,dt.$$

6. If m is a function of x, and $\mathfrak{A}$ is a vector which depends on x, and if each has a derivative, show that $m\mathfrak{A}$ has a derivative, and that

$$\frac{d(m\mathfrak{A})}{dx} = \frac{dm}{dx}\mathfrak{A} + m\frac{d\mathfrak{A}}{dx}.$$

7. If a point P move in any manner in space, its coordinates being given by the equations

$$x = f(t), \qquad y = \phi(t), \qquad z = \psi(t),$$

where f, ϕ, ψ are continuous functions of the time, having continuous derivatives, and if

$$\mathfrak{r} = x\mathfrak{i} + y\mathfrak{j} + z\mathfrak{k},$$

show that the vector velocity of P is represented by

$$\dot{\mathfrak{r}} = \frac{d\mathfrak{r}}{dt}.$$

8. If f, ϕ, ψ have continuous second derivatives, show that the vector acceleration of P is given by

$$\ddot{\mathfrak{r}} = \frac{d^2\mathfrak{r}}{dt^2}.$$

9. Show that the plane determined by the vectors $\dot{\mathfrak{r}}$ and $\ddot{\mathfrak{r}}$ drawn from P (on the assumption that neither is a nul vector) is the osculating plane. Thus the vector acceleration always lies in the osculating plane.

* Two vectors are said to be *collinear*, if there is a line in space to which they are both parallel.

2. The Scalar Product. Beside vector addition (§ 1) there are two other laws * whereby two given vectors determine a new quantity, — namely, the *scalar product* and the *vector product*. The first of these laws is as follows. Let $\mathfrak{A}$ and $\mathfrak{B}$ be any two vectors, and let the angle between them ** be ϵ. Then the number $|\mathfrak{A}| \cdot |\mathfrak{B}| \cdot \cos\epsilon$ is defined as their *scalar product*, and is written $S\mathfrak{A}\mathfrak{B}$:

$$S\mathfrak{A}\mathfrak{B} = |\mathfrak{A}| \cdot |\mathfrak{B}| \cdot \cos\epsilon$$

or $$= AB\cos\epsilon, \qquad A = |\mathfrak{A}|, \quad B = |\mathfrak{B}|.$$

If one of the vectors $\mathfrak{A}$ or $\mathfrak{B}$ is a nul vector, the scalar product is defined as 0.

Thus the scalar product of two vectors is not a vector at all, but is an ordinary number. Such numbers were called by Hamilton, the founder of quaternions, *scalars* because they can be represented by points on a scale, or graduated ruler.

The scalar product vanishes (*i*) when one of the factors vanishes; (*ii*) when the given vectors are at right angles with each other.

The scalar product can be interpreted (or, if one will, *defined*) as follows. Let each of the given vectors be projected on a line parallel to or coincident with one of them. Then the product of these projections, each taken with its proper sign, is equal to the scalar product in question.

The *commutative law* holds for scalar multiplication (as it is also called):

$$S\mathfrak{A}\mathfrak{B} = S\mathfrak{B}\mathfrak{A}.$$

But the *associative law* has no meaning, since the definition of scalar multiplication applies only to two vectors, not to a scalar and a vector.

The *distributive law*, on the other hand, is true here:

$$S\mathfrak{A}(\mathfrak{B} + \mathfrak{C}) = S\mathfrak{A}\mathfrak{B} + S\mathfrak{A}\mathfrak{C}.$$

The proof can be given by projecting the vectors along the line of $\mathfrak{A}$ and observing that

$$\text{Proj}\,(\mathfrak{B} + \mathfrak{C}) = \text{Proj}\,\mathfrak{B} + \text{Proj}\,\mathfrak{C}.$$

On applying the interpretation of the definition mentioned above, the truth of the theorem becomes apparent.

* The word *law* is here used in the sense of *definition*, not *theorem*.

** By this is meant the angle formed by two rays which emanate from a point and have respectively the direction and sense of the given vectors. This angle is an unsigned quantity and is taken between 0° and 180°, both inclusive.

Cartesian Form of the Scalar Product. Let

$$\mathfrak{A} = A_1\mathfrak{i} + A_2\mathfrak{j} + A_3\mathfrak{k},$$
$$\mathfrak{B} = B_1\mathfrak{i} + B_2\mathfrak{j} + B_3\mathfrak{k}.$$

Since
$$\begin{cases} S\,\mathfrak{i}\mathfrak{i} = 1, & S\,\mathfrak{j}\mathfrak{j} = 1, & S\,\mathfrak{k}\mathfrak{k} = 1, \\ S\,\mathfrak{j}\mathfrak{k} = 0, & S\,\mathfrak{k}\mathfrak{i} = 0, & S\,\mathfrak{i}\mathfrak{j} = 0; \end{cases}$$

and since, moreover, scalar multiplication is distributive, it follows at once that
$$S\,\mathfrak{A}\mathfrak{B} = A_1B_1 + A_2B_2 + A_3B_3.$$

The formula follows also from the definition, provided neither $\mathfrak{A}$ nor $\mathfrak{B}$ is a nul vector, since

$$\cos\epsilon = \frac{A_1B_1 + A_2B_2 + A_3B_3}{|\mathfrak{A}|\cdot|\mathfrak{B}|}.$$

EXERCISES

1. Show that
$$S\,\mathfrak{A}\mathfrak{A} = |\mathfrak{A}|^2.$$

2. If each of the vectors $\mathfrak{A}$ and $\mathfrak{B}$ has a derivative with respect to x, show that the scalar product has a derivative, and that

$$\frac{dS\,\mathfrak{A}\mathfrak{B}}{dx} = S\frac{d\mathfrak{A}}{dx}\mathfrak{B} + S\,\mathfrak{A}\frac{d\mathfrak{B}}{dx}.$$

3. If $\mathfrak{a}$ is a unit vector, $|\mathfrak{a}| = 1$, show that
$$S\,\mathfrak{a}\mathfrak{a}' = 0, \qquad \text{where} \qquad \mathfrak{a}' = \frac{d\mathfrak{a}}{dx}.$$

4. In the case of motion in a plane we have, on introducing polar coordinates,
$$\mathfrak{r} = \mathfrak{i}\,r\cos\phi + \mathfrak{j}\,r\sin\phi,$$

where r and ϕ are functions of the time.

Deduce the usual formulas for the components of the velocity and the acceleration along the radius vector (produced) and perpendicular to it, by considering the scalar products, $S\dot{\mathfrak{r}}\mathfrak{a}_r$, $S\dot{\mathfrak{r}}\mathfrak{a}_\phi$, $S\ddot{\mathfrak{r}}\mathfrak{a}_r$, $S\ddot{\mathfrak{r}}\mathfrak{a}_\phi$, where $\mathfrak{a}_r$ and $\mathfrak{a}_\phi$ denote unit vectors along $\mathfrak{r}$ and perpendicular to it respectively, and the dots denote time-derivatives.

Ans. $$v_r = \frac{dr}{dt}, \quad v_\phi = r\frac{d\phi}{dt}; \quad \alpha_r = \frac{d^2r}{dt^2} - r\frac{d\phi^2}{dt^2}, \quad \alpha_\phi = \frac{1}{r}\frac{d}{dt}\left(r^2\frac{d\phi}{dt}\right).$$

5. Show that the equation of a plane which (*i*) goes through a given point, A, and (*ii*) is perpendicular to a given vector $\mathfrak{A}$ can be written as follows. Let O be a fixed point in space, and let $\mathfrak{r}_0$ be

the vector drawn from O to A. Let P be any point of the plane, and let $\mathfrak{r}$ be the vector drawn from O to P. Then

$$S\mathfrak{A}(\mathfrak{r}-\mathfrak{r}_0)=0.$$

6. Show that the law of multiplication of determinants can be written in the form:

$$\begin{vmatrix} A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \\ C_1 & C_2 & C_3 \end{vmatrix} \cdot \begin{vmatrix} A_1' & A_2' & A_3' \\ B_1' & B_2' & B_3' \\ C_1' & C_2' & C_3' \end{vmatrix} = \begin{vmatrix} S\mathfrak{A}\mathfrak{A}' & S\mathfrak{A}\mathfrak{B}' & S\mathfrak{A}\mathfrak{C}' \\ S\mathfrak{B}\mathfrak{A}' & S\mathfrak{B}\mathfrak{B}' & S\mathfrak{B}\mathfrak{C}' \\ S\mathfrak{C}\mathfrak{A}' & S\mathfrak{C}\mathfrak{B}' & S\mathfrak{C}\mathfrak{C}' \end{vmatrix}.$$

3. The Vector Product. Let $\mathfrak{A}$ and $\mathfrak{B}$ be any two vectors. Construct them with the same initial point, and complete the parallelogramme, of which they form two sides. Then the *vector product* of $\mathfrak{A}$ and $\mathfrak{B}$, written $V\mathfrak{A}\mathfrak{B}$, is defined as a vector drawn at right angles to the plane of the parallelogramme, of length equal to the area of the parallelogramme, and in a sense such that $\mathfrak{A}$, $\mathfrak{B}$, and $V\mathfrak{A}\mathfrak{B}$ shall always form a right-handed system, or else always form a left-handed system. Thus, when a Cartesian system of coordinates is introduced, we shall always have:

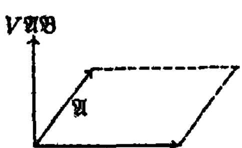


FIG. 70

(1) $$V\mathfrak{j}\mathfrak{k}=\mathfrak{i}, \qquad V\mathfrak{k}\mathfrak{i}=\mathfrak{j}, \qquad V\mathfrak{i}\mathfrak{j}=\mathfrak{k}.$$

If one of the vectors is a nul vector, or if the vectors are collinear, the vector product is defined as a nul vector.

The absolute value of the vector product is given as follows:

(2) $$|V\mathfrak{A}\mathfrak{B}| = AB\sin\epsilon,$$

where ϵ is the angle between the vectors.

It is clear that two vectors, $\mathfrak{A}$ and $\mathfrak{B}$, drawn from the same point, can be replaced in a great variety of ways by two other vectors, $\mathfrak{A}'$ and $\mathfrak{B}'$, lying in the same plane, without altering the value of the vector product. Thus in the figure

(3) $$V\mathfrak{A}\mathfrak{B}=V\mathfrak{A}_1\mathfrak{B}=V\mathfrak{A}_2\mathfrak{B}=\cdots, \quad\text{or}\quad V(\mathfrak{A}+m\mathfrak{B})\mathfrak{B}=V\mathfrak{A}\mathfrak{B}.$$

Moreover,

(4) $$V\mathfrak{A}\mathfrak{B}=V(m\mathfrak{A})\left(\frac{\mathfrak{B}}{m}\right),$$

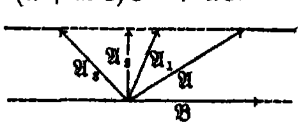


FIG. 71

where m is any number not 0. Hence one of the factors in a product can be replaced by a unit vector having the same direction and sense, and the direction and sense of the other factor will not thereby be disturbed.

The vector product vanishes (*i*) when one of the factors vanishes; (*ii*) when the two given vectors are collinear.

The *commutative* law does not hold; but we have

(5) $$V\mathfrak{A}\mathfrak{B} = -V\mathfrak{B}\mathfrak{A}.$$

The *associative* law does not hold; for

$$V\{\mathfrak{i}V\mathfrak{i}\mathfrak{j}\} = V\mathfrak{i}\mathfrak{k} = -\mathfrak{j}; \qquad \text{but} \qquad V\{V\mathfrak{i}\mathfrak{i}\}\mathfrak{i} = 0.$$

On the other hand, the *distributive* law is true:

(6) $$\begin{cases} V\mathfrak{A}(\mathfrak{B}+\mathfrak{C}) = V\mathfrak{A}\mathfrak{B} + V\mathfrak{A}\mathfrak{C}, \\ V(\mathfrak{B}+\mathfrak{C})\mathfrak{A} = V\mathfrak{B}\mathfrak{A} + V\mathfrak{C}\mathfrak{A}. \end{cases}$$

To prove this law, construct $\mathfrak{A}$, $\mathfrak{B}$, $\mathfrak{C}$, and $\mathfrak{D} = \mathfrak{B} + \mathfrak{C}$ from the same initial point, O. The terminal points of the last three vectors form with O the vertices of a parallelogram, $OBDC$. Let this parallelogram be projected on a plane perpendicular to $\mathfrak{A}$. The new quadrilateral, $O'B'D'C'$, will also be a parallelogram. Hence if we denote the vectors represented by its sides and diagonal by

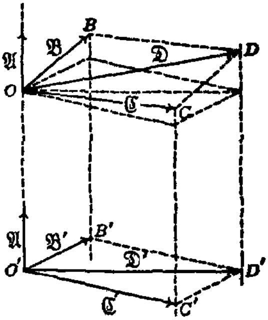


Fig. 72

$(O'B') = \mathfrak{B}'$, $\quad (O'C') = \mathfrak{C}'$, $\quad (O'D') = \mathfrak{D}'$,

we shall have $$\mathfrak{D}' = \mathfrak{B}' + \mathfrak{C}'.$$

Now, it is clear from the geometric construction that

(7) $$V\mathfrak{A}\mathfrak{B} = V\mathfrak{A}\mathfrak{B}', \qquad V\mathfrak{A}\mathfrak{C} = V\mathfrak{A}\mathfrak{C}', \qquad V\mathfrak{A}\mathfrak{D} = V\mathfrak{A}\mathfrak{D}'.$$

Moreover, as we will presently show,

(8) $$V\mathfrak{A}\mathfrak{B}' + V\mathfrak{A}\mathfrak{C}' = V\mathfrak{A}\mathfrak{D}'.$$

For, let the parallelogram $O'B'D'C'$ be rotated about the line of $\mathfrak{A}$ through 90° so that $O'B'$, in its new position, will lie along $V\mathfrak{A}\mathfrak{B}'$. Then $O'C'$ comes to lie along $V\mathfrak{A}\mathfrak{C}'$, and $O'D'$ to lie along $V\mathfrak{A}\mathfrak{D}'$. Let the new positions of B', D', C' be denoted respectively by B'', D'', C''. Then

$$(O'B'') = \frac{V\mathfrak{A}\mathfrak{B}'}{|\mathfrak{A}|}, \qquad (O'C'') = \frac{V\mathfrak{A}\mathfrak{C}'}{|\mathfrak{A}|}, \qquad (O'D'') = \frac{V\mathfrak{A}\mathfrak{D}'}{|\mathfrak{A}|}.$$

But $$(O'B'') + (O'C'') = (O'D'').$$

From (8) and (7) we now infer the truth of the theorem.

It has been tacitly assumed that no one of the vectors $\mathfrak{A}$, $\mathfrak{B}$, $\mathfrak{C}$, $\mathfrak{D}$ is a nul vector, and that neither $\mathfrak{B}$ nor $\mathfrak{C}$ is collinear with $\mathfrak{A}$. In each of these excepted cases the truth of the theorem is at once evident.

The second form of the distributive law follows from the first form by virtue of the relation (5).

Cartesian Form of the Vector Product. Let

$$\mathfrak{A} = A_1\,\mathfrak{i} + A_2\,\mathfrak{j} + A_3\,\mathfrak{k},$$
$$\mathfrak{B} = B_1\,\mathfrak{i} + B_2\,\mathfrak{j} + B_3\,\mathfrak{k}$$

be any two vectors. From the distributive law and the relations (1) we infer that

$$V\mathfrak{A}\mathfrak{B} = (A_2B_3 - A_3B_2)\,\mathfrak{i} + (A_3B_1 - A_1B_3)\,\mathfrak{j} + (A_1B_2 - A_2B_1)\,\mathfrak{k}.$$

This result is also expressed in the form

$$(9) \qquad V\mathfrak{A}\mathfrak{B} = \begin{vmatrix} \mathfrak{i} & \mathfrak{j} & \mathfrak{k} \\ A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \end{vmatrix}.$$

Application. Let a rigid body be rotating about an axis with angular velocity ω, and vector angular velocity (ω). To find the velocity, v, and the vector velocity, $\mathfrak{v} = (v)$, of a point P of the body.

Let h be the length of the perpendicular dropped from P on the axis. (The object of drawing in the frame of reference, or *Achsenkreuz*, is to enable the reader more easily to visualize the space figure.) Then

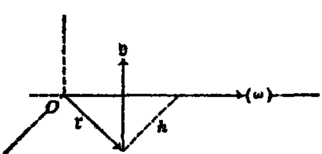


Fig. 73

$$v = h\omega.$$

The direction of the motion of P is at right angles to the plane determined by the axis and P. Let O be an arbitrary point of the axis, and let $\mathfrak{r} = (OP)$. Then

$$\mathfrak{v} = V(\omega)\,\mathfrak{r}.$$

The Cartesian form of this result is as follows. Let the coordinates of O and P be respectively (a, b, c) and (x, y, z); and let the components of (ω) along the axes be denoted by ω_1, ω_2, ω_3. Then

$$\mathfrak{v} = [(z-c)\,\omega_2 - (y-b)\,\omega_3]\,\mathfrak{i} + [(x-a)\,\omega_3 - (z-c)\,\omega_1]\,\mathfrak{j}$$
$$+ [(y-b)\,\omega_1 - (x-a)\,\omega_2]\,\mathfrak{k},$$

or

$$v_1 = (z - c)\,\omega_2 - (y - b)\,\omega_3,$$
$$v_2 = (x - a)\,\omega_3 - (z - c)\,\omega_1,$$
$$v_3 = (y - b)\,\omega_1 - (x - a)\,\omega_2.$$

EXERCISES

1. If $\mathfrak{A}$, $\mathfrak{B}$, and $\mathfrak{C}$ are three non-complanar vectors with the same initial point, the volume of the parallelepiped determined by them is equal numerically to

$$S\mathfrak{A}V\mathfrak{B}\mathfrak{C}.$$

2. Show that a necessary and sufficient condition that three vectors, $\mathfrak{A}$, $\mathfrak{B}$, and $\mathfrak{C}$, be complanar is that

$$S\mathfrak{A}V\mathfrak{B}\mathfrak{C} = 0.$$

3. If $\mathfrak{A}$ and $\mathfrak{B}$ be two vectors, each of which admits a derivative with respect to x, show that their vector product admits a derivative with respect to x, and that

$$\frac{d}{dx}V\mathfrak{A}\mathfrak{B} = V\frac{d\mathfrak{A}}{dx}\mathfrak{B} + V\mathfrak{A}\,\frac{d\mathfrak{B}}{dx}.$$

4. Prove that the equation of a plane which (i) passes through a given point, A, and (ii) is parallel to each of two non-collinear vectors, $\mathfrak{A}$ and $\mathfrak{B}$, can be written in the form:

$$S(\mathfrak{r} - \mathfrak{r}_0)\,V\mathfrak{A}\mathfrak{B} = 0,$$

where $\mathfrak{r}_0$, $\mathfrak{r}$ are vectors drawn from a fixed point, O, in space to A and an arbitrary point, P, of the plane

5. Show that the equation of the osculating plane of the space curve, § 1, Ex. 7 is

$$S(\mathfrak{R} - \mathfrak{r})\,V\dot{\mathfrak{r}}\ddot{\mathfrak{r}} = 0,$$

where $\mathfrak{R}$ is the vector drawn from the origin to an arbitrary point of the plane.

4. Rotation of the Axes; Direction Cosines. The formulas whereby we pass from one system of Cartesian axes to a second having the same origin (both systems being right-handed, or else both left-handed) are important, not only in analytic geometry of three dimensions, but also in mathematical physics. Let the direction cosines of the axis of x', referred to the (x, y, z)-axes, be l_1, m_1, n_1; those of y' and z' being formed by advancing the subscripts. These definitions are succinctly set forth by the following tables, the second one applying to unit vectors laid off along the axes:

(1)

	x'	y'	z'
x	l_1	l_2	l_3
y	m_1	m_2	m_3
z	n_1	n_2	n_3

	$\mathfrak{i}'$	$\mathfrak{j}'$	$\mathfrak{k}'$
$\mathfrak{i}$	l_1	l_2	l_3
$\mathfrak{j}$	m_1	m_2	m_3
$\mathfrak{k}$	n_1	n_2	n_3

Thus we have

$$(2)\quad \begin{cases} x = l_1x' + l_2y' + l_3z' \\ y = m_1x' + m_2y' + m_3z' \\ z = n_1x' + n_2y' + n_3z' \end{cases} \qquad \begin{cases} \mathfrak{i} = l_1\mathfrak{i}' + l_2\mathfrak{j}' + l_3\mathfrak{k}' \\ \mathfrak{j} = m_1\mathfrak{i}' + m_2\mathfrak{j}' + m_3\mathfrak{k}' \\ \mathfrak{k} = n_1\mathfrak{i}' + n_2\mathfrak{j}' + n_3\mathfrak{k}' \end{cases}$$

with analogous formulas for x', y', z' (or $\mathfrak{i}', \mathfrak{j}', \mathfrak{k}'$) in terms of x, y, z (or $\mathfrak{i}, \mathfrak{j}, \mathfrak{k}$) obtained by reading down the table, instead of across.

Formulas (2), in either form, rest for their proof on that most important principle, which we meet early in trigonometry and use so often in analytic geometry, that *if two broken lines in space have the same extremities, the sum of the projections of the one line on an arbitrary line in space is equal to the sum of the projections of the other line on that same line.*

From the fact that the three letters which stand in any row or in any column are the direction cosines of a line we see that

$$(3)\quad \begin{cases} l_1^2 + l_2^2 + l_3^2 = 1 \\ m_1^2 + m_2^2 + m_3^2 = 1 \\ n_1^2 + n_2^2 + n_3^2 = 1 \end{cases} \qquad \begin{cases} l_1^2 + m_1^2 + n_1^2 = 1 \\ l_2^2 + m_2^2 + n_2^2 = 1 \\ l_3^2 + m_3^2 + n_3^2 = 1 \end{cases}$$

From the fact that each system of coordinates is an orthogonal system we have that

$$(4)\quad \begin{cases} l_2l_3 + m_2m_3 + n_2n_3 = 0 \\ l_3l_1 + m_3m_1 + n_3n_1 = 0 \\ l_1l_2 + m_1m_2 + n_1n_2 = 0 \end{cases} \qquad \begin{cases} m_1n_1 + m_2n_2 + m_3n_3 = 0 \\ n_1l_1 + n_2l_2 + n_3l_3 = 0 \\ l_1m_1 + l_2m_2 + l_3m_3 = 0 \end{cases}$$

Thus far vector analysis has not been needed, since the relations (3) and (4) follow directly from the definitions. There is, however, a further system of relations which is established with great ease by means of the vector product. Starting with the relation

$$V\mathfrak{i}\mathfrak{j} = \mathfrak{k},$$

compute each side in terms of $\mathfrak{i}', \mathfrak{j}', \mathfrak{k}'$.

$$V\mathfrak{i}\mathfrak{j} = V(l_1\mathfrak{i}' + l_2\mathfrak{j}' + l_3\mathfrak{k}')(m_1\mathfrak{i}' + m_2\mathfrak{j}' + m_3\mathfrak{k}')$$
$$= (l_2 m_3 - l_3 m_2)\,\mathfrak{i}' + (l_3 m_1 - l_1 m_3)\,\mathfrak{j}' + (l_1 m_2 - l_2 m_1)\mathfrak{k}',$$
$$\mathfrak{k} = n_1\mathfrak{i}' + n_2\mathfrak{j}' + n_3\mathfrak{k}'.$$

On equating coefficients we obtain the third triple in the following system of relations, the other two being found from the remaining equations (1), § 3.

$$(5)\quad \begin{cases} l_1 = m_2 n_3 - m_3 n_2 \\ l_2 = m_3 n_1 - m_1 n_3 \\ l_3 = m_1 n_2 - m_2 n_1 \end{cases} \quad \begin{cases} m_1 = n_2 l_3 - n_3 l_2 \\ m_2 = n_3 l_1 - n_1 l_3 \\ m_3 = n_1 l_2 - n_2 l_1 \end{cases} \quad \begin{cases} n_1 = l_2 m_3 - l_3 m_2 \\ n_2 = l_3 m_1 - l_1 m_3 \\ n_3 = l_1 m_2 - l_2 m_1 \end{cases}$$

A further system of relations is obtained by writing equations (1) of § 3 for $\mathfrak{i}'$, $\mathfrak{j}'$, $\mathfrak{k}'$ and proceeding in a similar manner. They are the following:

$$(6)\quad \begin{cases} l_1 = m_2 n_3 - m_3 n_2 \\ m_1 = n_2 l_3 - n_3 l_2 \\ n_1 = l_2 m_3 - l_3 m_2 \end{cases} \quad \begin{cases} l_2 = m_3 n_1 - m_1 n_3 \\ m_2 = n_3 l_1 - n_1 l_3 \\ n_2 = l_3 m_1 - l_1 m_3 \end{cases} \quad \begin{cases} l_3 = m_1 n_2 - m_2 n \\ m_3 = n_1 l_2 - n_2 l_1 \\ n_3 = l_1 m_2 - l_2 m_1 \end{cases}$$

Finally, it is now easy to show that

$$(7)\qquad \begin{vmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \\ n_1 & n_2 & n_3 \end{vmatrix} = 1.$$

5. Invariants. We have given a definition of the scalar product of two vectors, which makes that product depend only on the vectors involved, and not on any system of coordinates. Then we evaluated it in terms of a Cartesian system of axes and found that

$$(1)\qquad S\mathfrak{A}\mathfrak{B} = A_1B_1 + A_2B_2 + A_3B_3.$$

Referred to a second system of Cartesian axes, obtained from the first by the transformation (2), § 4, the scalar product has the value

$$(2)\qquad S\mathfrak{A}\mathfrak{B} = A_1'B_1' + A_2'B_2' + A_3'B_3'.$$

Hence

$$(3)\qquad A_1'B_1' + A_2'B_2' + A_3'B_3' = A_1B_1 + A_2B_2 + A_3B_3.$$

Such an expression is called an *invariant** with respect to the transformation because, when formed for the A''s and B''s, it has the same value as when formed for the A's and B's.

* For a general discussion of the idea of *invariant* cf. Bôcher, *Algebra*, Chap. VII.

The proof given above that the expression

$$(4) \qquad A_1B_1 + A_2B_2 + A_3B_3$$

is an invariant was geometric. By virtue of the relation (1), this expression was identified with a quantity, $S\mathfrak{A}\mathfrak{B}$, which from its very definition is invariant. It would be possible, however, to give a direct algebraic proof by computing the A's and B's in terms of the A''s and B''s, substituting these values in (4), and reducing. The student should carry through this proof.

The Vector Product. The situation is similar with reference to the vector product. This vector, like the scalar product, depends only on the two given vectors, not on any system of coordinates. Its value in terms of a Cartesian system of axes has been found to be:

$$(5) \qquad V\mathfrak{A}\mathfrak{B} = \begin{vmatrix} \mathfrak{i} & \mathfrak{j} & \mathfrak{k} \\ A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \end{vmatrix}.$$

Hence the expression which stands on the right, when formed for a new system of axes as given by (2), § 4, each letter that enters now being primed, must have the same value, since it represents the same vector.

Again, the proof is geometric. An analytic proof can, however, be given directly by computing the A's, B's, $\mathfrak{i}$, $\mathfrak{j}$, $\mathfrak{k}$ in terms of the A''s, B''s, $\mathfrak{i}'$, $\mathfrak{j}'$, $\mathfrak{k}'$, substituting these values in the determinant, and reducing. The law of the multiplication of determinants, and relation (7) of § 4, are here involved. The student should carry through the details.

It is not merely to satisfy an æsthetic desire for completeness or a moral desire for truth, that we have given the above discussion of the two aspects of each of the invariants — the geometric definition and the analytic expression — but rather to provide ourselves with the means of dealing with the *symbolic vectors* which enter in the next paragraph. For there, the geometric definition ceases to exist and we are obliged to fall back on the analytic form.

6. Symbolic Vectors. Curl. Let u be a function of x, y, z, continuous together with its derivatives of the first order, and let it be carried over by the linear transformation (2), § 4, into a function of x', y', z'. Then

$$(1)\qquad \begin{cases} \dfrac{\partial u}{\partial x'} = l_1 \dfrac{\partial u}{\partial x} + m_1 \dfrac{\partial u}{\partial y} + n_1 \dfrac{\partial u}{\partial z}, \\[2mm] \dfrac{\partial u}{\partial y'} = l_2 \dfrac{\partial u}{\partial x} + m_2 \dfrac{\partial u}{\partial y} + n_2 \dfrac{\partial u}{\partial z}, \\[2mm] \dfrac{\partial u}{\partial z'} = l_3 \dfrac{\partial u}{\partial x} + m_3 \dfrac{\partial u}{\partial y} + n_3 \dfrac{\partial u}{\partial z}. \end{cases}$$

Thus it appears that, when x, y, z are subjected to the linear transformation (2), § 4, the partial derivatives of u with respect to x, y, z are transformed by precisely the same law.

If we expunge from these equations the letter u altogether and interpret the marks

$$\frac{\partial}{\partial x},\ \frac{\partial}{\partial y},\ \frac{\partial}{\partial z} \qquad \text{and} \qquad \frac{\partial}{\partial x'},\ \frac{\partial}{\partial y'},\ \frac{\partial}{\partial z'}$$

as ordinary numbers, then equations (1), modified as prescribed, represent the same linear transformation (2), § 4, performed on these variables. The equation

$$(2)\qquad \nabla = \mathfrak{i}\frac{\partial}{\partial x} + \mathfrak{j}\frac{\partial}{\partial y} + \mathfrak{k}\frac{\partial}{\partial z}$$

represents an ordinary vector, the expression for which, after the transformation has been performed, is

$$\nabla = \mathfrak{i}'\frac{\partial}{\partial x'} + \mathfrak{j}'\frac{\partial}{\partial y'} + \mathfrak{k}'\frac{\partial}{\partial z'}.$$

Thus if

$$\mathfrak{A} = A\mathfrak{i} + B\mathfrak{j} + C\mathfrak{k}$$

be an ordinary vector, we have,

$$(3)\qquad S\nabla\mathfrak{A} = \frac{\partial}{\partial x}A + \frac{\partial}{\partial y}B + \frac{\partial}{\partial z}C,$$

and this quantity, as we saw in § 5, is invariant with respect to any transformation (2), § 4.

We now proceed to show that, if we interpret the symbols on the right of this last equation as meaning differentiations, the same result is true, or

$$(4)\qquad \frac{\partial A'}{\partial x'} + \frac{\partial B'}{\partial y'} + \frac{\partial C'}{\partial z'} = \frac{\partial A}{\partial x} + \frac{\partial B}{\partial y} + \frac{\partial C}{\partial z},$$

where A, B, C are any three functions which, together with their first partial derivatives, are continuous. We have:

$$\frac{\partial A'}{\partial x'} = l_1\frac{\partial A'}{\partial x} + m_1\frac{\partial A'}{\partial y} + n_1\frac{\partial A'}{\partial z}, \qquad A' = l_1A + m_1B + n_1C,$$

with similar formulas for $\partial B'/\partial y'$ and $\partial C''/\partial z'$. Thus it appears that the actual computation of $\partial A'/\partial x'$ is identical *in form* with the evaluation of the product *

$$\left(l_1\frac{\partial}{\partial x}+m_1\frac{\partial}{\partial y}+n_1\frac{\partial}{\partial z}\right)\left(l_1A+m_1B+n_1C\right),$$

and the subsequent interpretation of such products as $\frac{\partial}{\partial x}A$, etc., as meaning differentiations.

It is in this sense that symbolic vectors are to be understood. They are not vectors in any geometrical sense, and the geometrical definitions of the scalar product and the vector product do not apply to them. They are vectors only in the sense of *algebraic form*, and their definition must in the nature of the case be *purely algebraic*. Obviously, the discussion in § 5 must underlie any real understanding of what is going on here. Those vector analysts who omit such a discussion put themselves into the class of people who justify the means by the end.

The invariant

(5) $$S\nabla\mathfrak{A}=\frac{\partial A}{\partial x}+\frac{\partial B}{\partial y}+\frac{\partial C}{\partial z}$$

is known as the *divergence* of the vector $\mathfrak{A}$ and is denoted by div $\mathfrak{A}$:.

(6) $$\text{div}\,\mathfrak{A}=S\nabla\mathfrak{A}=\frac{\partial A}{\partial x}+\frac{\partial B}{\partial y}+\frac{\partial C}{\partial z}.$$

Curl. Consider the vector product

(7) $$V\nabla\mathfrak{A}=\begin{vmatrix}\mathfrak{i} & \mathfrak{j} & \mathfrak{k}\\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z}\\ A & B & C\end{vmatrix}.$$

This vector is invariant for geometric reasons when we interpret ∇ as an ordinary vector. It is invariant for algebraic reasons when we interpret ∇ as a symbolic vector. With the latter interpretation it is known as the *curl* of the given vector, $\mathfrak{A}$:

* It is obviously immaterial whether we prove that the right-hand side of (4) can be transformed into the left-hand side, or the left-hand side into the right-hand.

(8) $\operatorname{curl} \mathfrak{A} = V\nabla\mathfrak{A} = \left(\frac{\partial C}{\partial y} - \frac{\partial B}{\partial z}\right)\mathfrak{i} + \left(\frac{\partial A}{\partial z} - \frac{\partial C}{\partial x}\right)\mathfrak{j} + \left(\frac{\partial B}{\partial x} - \frac{\partial A}{\partial y}\right)\mathfrak{k}.$

The Invariant $S\nabla\nabla u = \Delta u$. It is possible to write Laplace's operator,

(9) $$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2},$$

in the form of a symbolic scalar product, and thus bring out its invariant character, for we see that

(10) $$S\nabla\nabla = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}.$$

It is for this reason that Laplace's operator is sometimes written as ∇^2 (read *triangle squared*). Thus

(11) $$\nabla u = \mathfrak{i}\frac{\partial u}{\partial x} + \mathfrak{j}\frac{\partial u}{\partial y} + \mathfrak{k}\frac{\partial u}{\partial z}$$

is a vector function — the so-called *gradient* of u, an invariant under the transformations (2), § 4 — whereas

(12) $$\nabla^2 u = S\nabla\nabla u = \Delta u$$

is a scalar, also invariant.

7. Green's Theorem and Stokes's Theorem in Vector Form. It is possible to write the main results of Chap. XI, §§ 9, 10, in vector form and thus bring out their invariant character. Thus Equation I, § 9, of that chapter appears as:

(1) $$\iiint_V S\nabla\mathfrak{A}\, dV = -\iint_S S\mathfrak{A}\nu\, dS,$$

where ν denotes a unit vector directed along the inner normal to the surface S.

In Ex. 2 at the close of that paragraph, the right-hand side of the equation is already in invariant form. The left-hand side can also be given such form by writing the integrands of the triple integrals respectively as $u\, S\nabla\nabla v$ and $S\nabla u\nabla v$.

Green's Theorem, as expressed in Ex. 3, is already in invariant form, since Laplace's operator is invariant.

Stokes's Theorem, § 10, I. of Chap. XI, now appears as

$$\int_S\int S(\operatorname{curl}\mathfrak{A})\nu\,dS = \int_C S\,\mathfrak{A}\,ds, \tag{2}$$

where $\mathfrak{A} = P\mathfrak{i} + Q\mathfrak{j} + R\mathfrak{k}$,

and thus its invariant character is established.

The condition that

$$\int_{(a,b,c)}^{(x,y,z)} P\,dx + Q\,dy + R\,dz$$

be independent of the path of integration can now be written in the vector form

$$\operatorname{curl}\mathfrak{A} = 0. \tag{3}$$

8. Curvature and Torsion of Twisted Curves. Frenet's Formulas. Let a space curve be given by the equations

$$x = f(s), \qquad y = \phi(s) \qquad z = \psi(s), \tag{1}$$

where s denotes the arc. Let

$$\mathfrak{r} = x\mathfrak{i} + y\mathfrak{j} + z\mathfrak{k}$$

be the vector drawn from the origin to a variable point of the curve. Then

$$\mathfrak{t} = \mathfrak{r}' = x'\mathfrak{i} + y'\mathfrak{j} + z'\mathfrak{k} \tag{2}$$

is a unit vector drawn in the sense of the positive tangent (the increasing s).

The vector $\mathfrak{r}''$ is normal to the curve, since $S\mathfrak{r}'\mathfrak{r}'' = 0$; § 2, Ex. 3. Moreover, it lies in the osculating plane; § 1, Ex. 9. In the case of a plane curve, its length is the curvature, and the definition is extended to twisted curves:

$$\kappa = \frac{1}{\rho} = |\mathfrak{r}''| = \sqrt{x''^2 + y''^2 + z''^2}. \tag{3}$$

Let $\mathfrak{n}$ be a unit vector taken along $\mathfrak{r}''$:

$$\mathfrak{n} = \rho\mathfrak{r}''. \tag{4}$$

This vector lies along the *principal normal*; Chap. VI, § 5. Finally, let

$$\nu = V\mathfrak{t}\mathfrak{n}. \tag{5}$$

Then ν is a third unit vector, and it lies along the *bi-normal*; i. e.

The three unit vectors,

(6) $$\mathfrak{t} = \mathfrak{r}', \qquad \mathfrak{n} = \rho\mathfrak{r}'', \qquad \nu = V\mathfrak{t}\mathfrak{n},$$

are oriented toward each other as the axes of x, y, z. It is convenient, moreover, to form the picture of these three vectors drawn from a fixed point, O, in space. If we think of P as describing the curve (1) with unit velocity, the terminal point T of $\mathfrak{t}$ will describe a path on the unit sphere about O with a velocity equal to the *curvature*, and the terminal point Q of ν will describe its path with a velocity equal numerically, as we shall presently see, to the *torsion*.

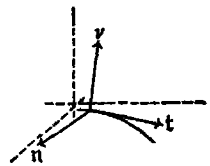


FIG. 74

The Derivatives, $\mathfrak{t}', \mathfrak{n}', \nu'$. From (2) and (4),

(7) $$\mathfrak{t}' = \frac{\mathfrak{n}}{\rho}.$$

The vector ν' is parallel to $\mathfrak{n}$, as seems plausible from the fact that the curve has contact of higher order with its osculating plane. A proof can be given as follows. Write (cf. § 1, Ex. 4)

$$\nu' = a\,\mathfrak{t} + b\,\mathfrak{n} + c\,\nu.$$

Since ν is a unit vector, $S\nu\nu' = 0$ by § 2, Ex. 3. Thus

$$0 = S\nu\nu' = aS\nu\mathfrak{t} + bS\nu\mathfrak{n} + cS\nu\nu.$$

Now ν is perpendicular to $\mathfrak{t}$ and $\mathfrak{n}$, so $S\nu\mathfrak{t} = 0$, $S\nu\mathfrak{n} = 0$. Moreover $S\nu\nu = 1$. Hence $c = 0$.

Next, differentiate the equation $S\mathfrak{t}\nu = 0$:

$$S\,\mathfrak{t}'\nu + S\,\mathfrak{t}\nu' = 0.$$

From (7) we see that $S\mathfrak{t}'\nu = 0$. Hence

$$0 = S\mathfrak{t}\nu' = aS\mathfrak{t}\mathfrak{t} + bS\mathfrak{t}\mathfrak{n} = a.$$

Thus $\nu' = b\mathfrak{n}$, and the theorem is proved.

The coefficient b is defined as the *torsion*; its reciprocal, as the *radius of torsion*, τ. Thus the torsion is a signed quantity (*i.e.* it may be either positive or negative); the curvature is unsigned.

We have, then, finally as the evaluation of ν':

(8) $$\nu' = \frac{\mathfrak{n}}{\tau}.$$

To compute the torsion, observe that it follows from (8) that

$$S\mathfrak{n}\nu' = \frac{1}{\tau}S\mathfrak{n}\mathfrak{n} = \frac{1}{\tau}.$$

Since $\nu = V\mathfrak{t}\mathfrak{n} = \rho V \mathfrak{r}'\mathfrak{r}''$,

we have: $\nu' = \rho V \mathfrak{r}'\mathfrak{r}''' + \rho' V \mathfrak{r}'\mathfrak{r}''$,

and so $S\mathfrak{n}\nu' = \rho S\mathfrak{r}''\nu' = \rho^2 S\mathfrak{r}'' V\mathfrak{r}'\mathfrak{r}'''$.

Hence finally

(9) $$\frac{1}{\tau} = -\frac{\begin{vmatrix} x' & y' & z' \\ x'' & y'' & z'' \\ x''' & y''' & z''' \end{vmatrix}}{x''^2 + y''^2 + z''^2}.$$

Computation of $\mathfrak{n}'$. It remains to compute $\mathfrak{n}'$. Since

$$\mathfrak{n} = V\nu\mathfrak{t},$$

we have: $$\mathfrak{n}' = V\nu\mathfrak{t}' + V\nu'\mathfrak{t}$$

$$= \frac{V\nu\mathfrak{n}}{\rho} + \frac{V\mathfrak{n}\mathfrak{t}}{\tau} = -\frac{\mathfrak{t}}{\rho} - \frac{\nu}{\tau}.$$

The results here established are known as

Frenet's Formulas : —

(10) $$\begin{cases} \dfrac{d\mathfrak{t}}{ds} = \ * \quad \dfrac{\mathfrak{n}}{\rho} \quad * \\[2ex] \dfrac{d\mathfrak{n}}{ds} = -\dfrac{\mathfrak{t}}{\rho} \quad * \quad -\dfrac{\nu}{\tau} \\[2ex] \dfrac{d\nu}{ds} = \ * \quad \dfrac{\mathfrak{n}}{\tau} \quad * \end{cases}$$

EXERCISES

The formulas for the curvature and the torsion, when x, y, z are expressed as functions of an arbitrary parameter t, are as follows:

(11) $$\frac{1}{\rho} = \frac{|V\dot{\mathfrak{r}}\ddot{\mathfrak{r}}|}{|\dot{\mathfrak{r}}|^3}, \qquad \frac{1}{\tau} = -\frac{S\dot{\mathfrak{r}}V\ddot{\mathfrak{r}}\dddot{\mathfrak{r}}}{|V\dot{\mathfrak{r}}\ddot{\mathfrak{r}}|^2},$$

where the dot denotes differentiation with respect to t.

The first of these is deduced by the aid of *Lagrange's Identity* relating to any four vectors:

(12) $$S\,V\mathfrak{A}\mathfrak{B}\,V\mathfrak{A}'\mathfrak{B}' = S\mathfrak{A}\mathfrak{A}'\ S\mathfrak{B}\mathfrak{B}' - S\mathfrak{A}\mathfrak{B}'\ S\mathfrak{A}'\mathfrak{B};$$

cf. Blaschke, *Differentialgeometrie*, vol. I, pp. 6 and 13.

Deduce these formulas with the help of the following suggestions

1. Show that $$\mathfrak{r}' = \frac{\dot{\mathfrak{r}}}{\sqrt{S\dot{\mathfrak{r}}\dot{\mathfrak{r}}}}.$$

2. Show that $$\mathfrak{r}'' = \frac{\ddot{\mathfrak{r}}}{S\dot{\mathfrak{r}}\dot{\mathfrak{r}}} - \frac{(S\dot{\mathfrak{r}}\ddot{\mathfrak{r}})\dot{\mathfrak{r}}}{(S\dot{\mathfrak{r}}\dot{\mathfrak{r}})^2}.$$

3. By means of Lagrange's Identity, prove that

$$S\dot{\mathfrak{r}}\dot{\mathfrak{r}}\,S\ddot{\mathfrak{r}}\ddot{\mathfrak{r}} - (S\dot{\mathfrak{r}}\ddot{\mathfrak{r}})^2 = |\,V\dot{\mathfrak{r}}\ddot{\mathfrak{r}}\,|^2.$$

Hence obtain Formula (11) for $1/\rho$.

4. Show that $$\mathfrak{r}''' = \frac{\dddot{\mathfrak{r}}}{(S\dot{\mathfrak{r}}\dot{\mathfrak{r}})^{3/2}} + a\,\dot{\mathfrak{r}} + b\,\ddot{\mathfrak{r}},$$

where a and b denote scalars which it is not required to compute.

Hence obtain equation (11) for $1/\tau$.

5. Prove Lagrange's Identity by direct computation, availing yourself of the simplifications rendered possible by the symmetry of the expressions.

6. Write out Equations (11) in ordinary form, not using any vector notation.

9. Notation. There are two aspects of vector analysis. One is formal, and has to do with the manipulation of algebraic identities. The other is geometric, and is chiefly concerned with the relation of the geometric concepts involved to facts of nature and the expression of these facts by equations between ordinary real quantities. The formal treatment necessarily lays much stress on notation and prizes highly the *product* idea. It writes the scalar product as

$$(\mathfrak{A}\mathfrak{B}) \qquad \text{or} \qquad \mathfrak{A}\cdot\mathfrak{B} \qquad \text{or} \qquad \mathfrak{A}\mathfrak{B}$$

and calls it, for no obvious reason, the *inner product*. And it writes the vector product as

$$[\mathfrak{A}\mathfrak{B}] \qquad \text{or} \qquad \mathfrak{A}\times\mathfrak{B} \qquad \text{or} \qquad \widehat{\mathfrak{A}\mathfrak{B}}$$

and calls it the *outer product*.

On the other hand, certain laws of physics involve the conception of vectors and the combinations treated in §§ 1–3, and so are best stated in vector form. It contributes to ease of comprehension to call a scalar a scalar and a vector a vector, and the old Hamiltonian notation of S and V, to which we have harked back, is self-explanatory.* When the vector expression of these laws has once been per-

* Hamilton's *scalar product* was the negative of that defined in § 2. The latter is the form which has been pretty generally accepted by later writers.

ceived, the next step is the translation into the form of three ordinary equations, involving the things with which ordinary analysis deals — integrals, derivatives, etc. — and the ultimate goal in the formulation of the physical law is the appreciation of the inner meaning of this last system of three ordinary equations — a meaning invariant of the choice of the coordinate axes.

So deeply impressed with the importance of this fact was the author of one of the greatest expositions of a physical subject ever written — Appell, in his *Mécanique rationelle* — that he dispenses altogether with the form of vector analysis and passes directly from the vector conceptions to the final result. Surely, this example may well give pause to those expositors who lean to a highly technical notation, the form of which is not geometrically suggestive, while algebraically it is so condensed as to require special training for its use. They are like the analysts who prove every theorem which has to do with a question of uniform convergence by direct use of ϵ's. The ϵ's are essential for establishing the basal theorems, just as the scalar and vector products are useful for establishing the basal theorems in certain branches of physics. These theorems once established — and their proofs involve but slight formal work, rather an appreciation of the inner content of the situation — it is the part of wisdom to work with these *results*, not *prove* them afresh every time one needs them. We do not compute our logarithms each time, before we use them. And so the general student can well dispense with these highly condensed notations, taking them up later if he has occasion to use them in some special and technical piece of work. When that time comes, he will find it most helpful to write his own *ABC* of vector analysis in terms of the notation he wishes to learn.

CHAPTER XIV

DIFFERENTIAL EQUATIONS

The student has already met a variety of differential equations in his study of the Calculus, and integrated them. The object of the present chapter is to systematize the methods which have hitherto been used, and to extend them. We shall, moreover, consider what the nature of the condition imposed on a function by a differential equation is and thus see how an approximate solution can be obtained.

1. **Ordinary Differential Equations.** An equation which connects a function, y, of a single independent variable, x, with its derivatives of the first n orders:

$$F\left(x, y, \frac{dy}{dx}, \frac{d^2y}{dx^2}, \cdots, \frac{d^ny}{dx^n}\right) = 0, \tag{1}$$

is called an *ordinary differential equation*, in distinction from a partial differential equation (cf. § 21), and its *order* is defined as n. If several functions, $y, z, \cdots$, are connected with one another and their derivatives by as many equations as there are functions, we have a *system* of ordinary differential equations. Thus the equations

$$\frac{dy}{dx} = F(x, y, z), \qquad \frac{dz}{dx} = \Phi(x, y, z) \tag{2}$$

form such a system.

By an *integral* (or *primitive*) of equation (1) is meant a function, $y = f(x)$, which satisfies that equation; *i.e.* if $f(x)$ be substituted for y in the left-hand member of (1):

$$F\left(x, f(x), \frac{df(x)}{dx}, \frac{d^2f(x)}{dx^2}, \cdots, \frac{d^nf(x)}{dx^n}\right),$$

this latter expression, which is a function of x alone, vanishes identically.

Each time that we have integrated a differential equation of the first order, we have found, as the most general solution, a function depending on one arbitrary constant. When we have integrated a differential equation of the second order, we have found a function depending on two arbitrary constants. And so we surmise that the most general function which satisfies equation (1) will depend on n arbitrary constants,

$$y = f(x, c_1, \cdots, c_n).$$

The guess is correct, and the precise formulation of the theorem will be given in §§ 15, 16.

As regards the processes admitted, we shall consider a function as known when it is given by a *quadrature, i.e.* when it is determined by an integral. This amounts to saying that, if

$$\frac{dy}{dx} = f(x),$$

where $f(x)$ is any function which is merely continuous in an interval $a < x < b$, and if x_0 is a point of this interval, the function

$$F(x) = \int_{x_0}^{x} f(x)\,dx$$

is considered as a known function (geometrically, it is the area under the curve), although it may not be possible to evaluate the integral in terms of the elementary functions; *i.e.* to express it in terms of rational functions, radicals, sines and cosines and their inverses, logarithms and exponentials, or as a combination of such functions. Example: the elliptic integral

$$\int_0^x \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}}.$$

We note that any *indefinite integral* which satisfies the above differential equation can be written as the *definite integral* whose value is $F(x)$, plus a constant, or

$$\int f(x)\,dx = \int_{x_0}^{x} f(x)\,dx + C.$$

Secondly, we regard a function as known when it is given by an *implicit equation.* Thus the equation

$$\log(x^2 + y^2) = \tan^{-1}\frac{y}{x} + C$$

defines y as a function of x, although we see no means of expressing y explicitly in terms of x by means of the elementary functions, and there is, moreover, no reason to suppose that such an expression is possible.

I. Equations of the First Order

2. Separation of Variables. Let the differential equation

$$\frac{dy}{dx} = f(x, y) \tag{1}$$

be given. It may be possible to *separate the variables, i.e.* so to transform the equation that all terms involving x appear on one side, and all terms involving y on the other side:

$$\frac{dx}{M(x)} = \frac{dy}{N(y)}.$$

The equation can then be integrated:

$$\int \frac{dx}{M(x)} = \int \frac{dy}{N(y)} + C,$$

where an arbitrary particular indefinite integral is chosen in each case.

The first example of the application of this method which we met appeared in the treatment of Simple Harmonic Motion, *Introduction to the Calculus*, Chap. XIII, § 6, p. 366:

$$\frac{dr}{dt} = -\sqrt{\frac{g}{R}}\sqrt{R^2 - r^2},$$

$$dt = -\sqrt{\frac{R}{g}}\frac{dr}{\sqrt{R^2 - r^2}},$$

$$t = -\sqrt{\frac{R}{g}}\int\frac{dr}{\sqrt{R^2 - r^2}} = \sqrt{\frac{R}{g}}\cos^{-1}\frac{r}{R} + C.$$

EXERCISES

Integrate each of the following differential equations: —

1. $\dfrac{dy}{dx} = 2xy.$ *Ans.* $y = Ce^{x^2}$.

2. $\dfrac{dy}{dx} = \dfrac{x}{y}.$ *Ans.* $x^2 - y^2 = C$.

3. $\dfrac{dy}{dx} = \dfrac{y}{x}.$ *Ans.* $y = Cx$.

4. $\sec x \cos^2 y \, dx = \cos x \sin y \, dy.$ *Ans.* $\sec y = \tan x + C$.

5. $$\sqrt{2ay-y^2}\csc x\,dx + y\tan x\,dy = 0.$$

Ans. $\csc x = a\cos^{-1}\dfrac{a-y}{a} - \sqrt{2ay-y^2} + C.$

6. $$x(3+x)\frac{dx}{dt} = t(2x+3).$$

7. $$x\sqrt{1-y^2}\,dx + y\sqrt{1-x^2}\,dy = 0.$$

8. $$(e^y+1)\cos x\,dx + e^y\sin x\,dy = 0.$$

3. Linear Equations. By a *linear differential equation of the first order* is meant an equation of the form *

(1) $$\frac{dy}{dx} + Py = Q,$$

where P and Q are given functions, depending on x, but not on y. Such an equation can always be integrated by means of the following device. Multiply through by the factor † $e^{\int P\,dx}$

(2) $$e^{\int P\,dx}\left(\frac{dy}{dx} + Py\right) = Qe^{\int P\,dx}.$$

The left-hand side is seen at once to be the derivative of the function $ye^{\int P\,dx}$. Hence (2) can be written in the form

(3) $$\frac{d}{dx}\left(ye^{\int P\,dx}\right) = Qe^{\int P\,dx},$$

and it remains merely to integrate each side of this last equation.

Example. Given
$$\frac{dy}{dx} + \frac{y}{x} = 4x^2.$$

Here, ‡ $$e^{\int P\,dx} = e^{\log x} = x.$$

On multiplying through by x we have:

$$x\frac{dy}{dx} + y = 4x^3$$

or $$\frac{d}{dx}(xy) = 4x^3.$$

* Extensions of the definition are found below in § 11 and § 27.

† Known as an *integrating factor*; cf. § 20.

‡ The student should review carefully Chap. VI, § 1, of the *Introduction to the Calculus*, for he will be expected, in the present chapter, to use freely the elementary properties of logarithms.

Hence $$xy = x^4 + C$$

or $$y = x^3 + \frac{C}{x}.$$

EXERCISES

Integrate each of the following differential equations :

1. $\frac{dy}{dx} - \frac{y}{x} = 4x^2.$ *Ans.* $y = 2x^3 + Cx$

2. $\frac{dy}{dx} - \frac{2xy}{x^2 + a^2} = 1.$ *Ans.* $y = \frac{x^2 + a^2}{a} \tan^{-1}\frac{x}{a} + C(x^2 + a^2).$

3. $\frac{dy}{dx} + y \cot x = \sec x.$ *Ans.* $y = (\log \sec x + C)\csc x.$

4. $x\frac{dy}{dx} + (1 + x)y = e^x.$ *Ans.* $y = \frac{e^x}{2x} + \frac{ce^{-x}}{x}.$

5. $\frac{dy}{dx} = x - y.$

6. $\cot\theta\frac{dr}{d\theta} = r + e^{-\theta}.$

7. $\frac{dy}{dx} + y = \cos x.$

8. $\frac{dy}{dx} - ay = b \sin x.$

9. Show that the differential equation

$$\frac{dz}{dx} + pz = \frac{q}{z^{n-1}}$$

can be reduced to a linear differential equation by the substitution

$$y = z^n.$$

10. Integrate the differential equation

$$\frac{ds}{dt} = as + \frac{b}{s}.$$

4. Homogeneous Equations. The differential equation

(1) $$\frac{dy}{dx} = \phi\left(\frac{y}{x}\right)$$

is sometimes called a *homogeneous equation,* because the right-hand side is homogeneous (of order 0) in x and y. It can be integrated by a quadrature. Introduce a new variable,*

$$v = \frac{y}{x}.$$

Then $\qquad y = vx, \qquad \dfrac{dy}{dx} = x\dfrac{dv}{dx} + v,$

and equation (1) becomes:

(2) $$x\frac{dv}{dx} + v = \phi(v).$$

Here, the variables are separable; we have

(3) $$\frac{dx}{x} = \frac{dv}{\phi(v) - v},$$

and it remains merely to integrate each side

Example. $$\frac{dy}{dx} = \frac{x+y}{x-y}.$$

$$v = \frac{y}{x}, \qquad x\frac{dv}{dx} + v = \frac{1+v}{1-v},$$

$$\frac{dx}{x} = \frac{1-v}{1+v^2}\,dv,$$

$$\log x = \tan^{-1}\frac{y}{x} - \tfrac{1}{2}\log(1+v^2) + C$$

$$= \tan^{-1}\frac{y}{x} - \log\sqrt{x^2+y^2} + \log x + C,$$

or

$$\tan^{-1}\frac{y}{x} = \log\sqrt{\frac{x^2+y^2}{a^2}}.$$

In polar coordinates this equation becomes:

$$\phi = \log\frac{r}{a}, \qquad r = ae^{\phi},$$

an equiangular spiral.

* The student will recall that a similar device was employed in algebra for the solution of two simultaneous quadratics of the form — the so-called *homogeneous case* —

$$a_1x^2 + b_1xy + c_1y^2 = d_1,$$
$$a_2x^2 + b_2xy + c_2y^2 = d_2.$$

EXERCISES

Solve the following differential equations:—

1. $$\frac{dy}{dx}=\frac{x-y}{x+y}.\qquad \textit{Ans. } x^2-2xy-y^2=C.$$

2. $$\frac{dy}{dx}=\frac{2x+3y}{x+4y}.\qquad \textit{Ans. } (x-y)^5(x+2y)=C.$$

3. $\dfrac{dy}{dx}=\dfrac{x+4y}{x+y}.$ 4. $\dfrac{dy}{dx}=\dfrac{x+y}{x}.$

5. $$\frac{dy}{dx}=\frac{Ax+By}{Cx+Dy}.$$ Treat all cases.

6. Show that the differential equation

$$\frac{dy}{dx}=\frac{x-y+3}{x+y-5}$$

can be reduced to the form considered in Question 1 by means of the transformation:

$$x=x'+h,\qquad y=y'+k$$

where h and k are determined by the equations:

$$h-k=-3,\qquad h+k=5.$$

7. Apply the method of Question 6 to the differential equation

$$\frac{dy}{dx}=\frac{Ax+By+L}{Cx+Dy+M},\qquad AD-BC\neq 0.$$

8. Solve the differential equation:

$$\frac{dy}{dx}=\frac{x+y}{x+y+1}.\qquad \textit{Ans. } 2(x-y)=\log(2x+2y+1)+C.$$

9. Show that the differential equation of Question 7, when $AD-BC=0$, can be solved by one of the substitutions

$$z=Ax+By,\qquad z=Cx+Dy,$$

provided all four coefficients $A, \cdots, D$ do not vanish.

10. Solve the differential equation

$$\frac{dy}{dx}=\frac{y}{x-c\sqrt{x^2+y^2}},$$

if $y=a$ when $x=0$.

$$\textit{Ans. } x=\frac{a}{2}\left\{\left(\frac{y}{a}\right)^{1-c}-\left(\frac{y}{a}\right)^{1+c}\right\}.$$

11. Show that the differential equation

$$\frac{dy}{dx} = f(ax + by), \quad a \neq 0, \quad b \neq 0,$$

can be solved by aid of the substitution

$$v = ax + by.$$

12. Integrate the equation:

$$\frac{dy}{dx} = \sin(2x + y).$$

13. Show that Equation (1) of the text admits the solution

$$y = ax,$$

in case the equation

$$\xi = \phi(\xi)$$

has a root,

$$\xi = a.$$

5. Equations of the Second Order with One Letter Absent. Consider the general differential equation of the second order,

$$\frac{d^2y}{dx^2} = F\left(x, y, \frac{dy}{dx}\right). \tag{1}$$

The function F will in general involve all three arguments, x, y, and dy/dx. If, however, one (or both) of the letters x and y is lacking, the equation can be reduced to one of the first order by means of the substitution

$$p = \frac{dy}{dx}. \tag{2}$$

If y is lacking, so that $F(x, y, p) = \phi(x, p)$, then (1) becomes

$$\frac{dp}{dx} = \phi(x, p). \tag{3}$$

If, however, it is x that fails, but y is present, so that

$$F(x, y, p) = \psi(y, p),$$

then, since

$$\frac{d^2y}{dx^2} = \frac{dy}{dx}\frac{dp}{dy} = p\frac{dp}{dy},$$

we have for (1) the equation

$$p\frac{dp}{dy} = \psi(y, p). \tag{4}$$

In each of these cases the solution of (1) has been reduced to the solution of a differential equation of the first order, followed by a quadrature.

A differential equation of the n-th order, which does not contain both the independent and the dependent variable explicitly, can be reduced in the same manner to one of lower order, and it remains, then, to integrate the latter and to perform a quadrature.

Example. We met this method for the first time in the chapter on *Mechanics*, when we integrated the differential equation

$$\frac{d^2s}{dt^2} = f(s)$$

by introducing the variable

$$v = \frac{ds}{dt},$$

thus reducing the original equation to the form

$$v\frac{dv}{ds} = f(s).$$

EXERCISES

1. Integrate the differential equation

$$(k^2 + a^2\cos^2\theta)\frac{d^2\theta}{dt^2} - a^2\sin\theta\cos\theta\left(\frac{d\theta}{dt}\right)^2 = -ag\sin\theta.$$

Ans. $(k^2 + a^2\cos^2\theta)\left(\frac{d\theta}{dt}\right)^2 = 2ag\cos\theta + C.$

2. Integrate:

$$y\frac{d^2x}{dy^2} = c\sqrt{1 + \left(\frac{dx}{dy}\right)^2},$$

if c is positive and $\neq 1$, and $y = 1$, $dx/dy = 0$, when $x = 0$.

Ans. $x = \frac{1}{2}\left\{\frac{y^{c+1}}{c+1} + \frac{y^{1-c}}{c-1} - \frac{2c}{c^2-1}\right\}.$

Applications

6. The Catenary. The catenary, as its name suggests, is the curve in which a chain hangs. Let us determine its equation.

The physical assumption is that of a material curve, homogeneous (*i.e.* of constant density, ρ), perfectly flexible, and inextensible, whose ends are fastened at two fixed points and which hangs at rest under the force of gravity.

Let the axis of x be horizontal, and at a distance c below the lowest point, A, of the curve, where c is a constant whose value we will assign later; and let the axis of y pass through A. Let $P:(x, y)$ be an arbitrary point of the curve, and

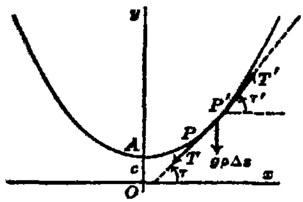


FIG. 75

$$P':(x+\Delta x, y+\Delta y),$$

a second point. Consider the arc PP' and *isolate this system; i.e.* consider this portion of matter and the forces which act on it. They are (*i*) the force of gravity, $g\rho\Delta s$, where ρ denotes the mass of one unit of length of the string;* (*ii*) the tension T at P, directed down † the tangent; (*iii*) the tension $T' = T + \Delta T$ at P', directed up † the tangent. Now, this arc PP' has been assumed flexible; but since it is in equilibrium, if it congeals and becomes a rigid body, it will obviously ‡ continue to be in equilibrium under the action of the above external forces, (*i*), (*ii*), and (*iii*). But we know the conditions under which a rigid body is in equilibrium under the action of a system of forces; in particular, it is necessary that they be such a system as would keep a particle at rest if they all acted at a point. Hence the algebraic sum of their components along an arbitrary direction must vanish.

Resolving, therefore, horizontally and vertically, we have:

(*a*) $$T'\cos\tau' = T\cos\tau;$$

(*b*) $$T'\sin\tau' = T\sin\tau + g\rho\,\Delta s.$$

From (*a*) we infer that

(1) $$T\cos\tau = T_0,$$

where T_0 denotes the tension at the lowest point, A. For, we may take P' at A, and then $T' = T_0$, $\tau' = 0$.

Equation (*b*) can be written in the form: §

* It is not important here that forces be measured in absolute units. If the student prefers, he may take w as the weight of one unit of length of the string, and then the force of gravity becomes $w\,\Delta s$.

† If P' lies below P, these directions will be reversed.

‡ This is, of course, the assumption of a new physical law, so compelling however, as to seem self-evident.

§ An equation analogous to (1) can be derived from (*b*), namely,

$$T\sin\tau = g\rho s,$$

where s is measured from O. Mechanically, this means that the vertical component of T at P is just equal to the weight of the arc in question.

$$(b') \qquad T' \sin \tau' - T \sin \tau = g\rho\, \Delta s.$$

The left-hand side of this equation is precisely the increment of the function $T \sin \tau$. Hence (b') becomes:

$$(b'') \qquad \Delta(T \sin \tau) = g\rho\, \Delta s.$$

Divide (b'') through by Δx and then let Δx approach the limit 0:

$$\lim_{\Delta x \doteq 0} \frac{\Delta(T \sin \tau)}{\Delta x} = g\rho \lim_{\Delta x \doteq 0} \frac{\Delta s}{\Delta x},$$

or *

$$(2) \qquad D_x(T \sin \tau) = g\rho\, D_x s.$$

Replacing T in (2) by its value, $T_0 \sec \tau$, from (1), we have:

$$T_0\, D_x(\tan \tau) = g\rho\, D_x s,$$

or, since $\tan \tau = D_x y$,

$$(3) \qquad D_x^2 y = \frac{1}{c} D_x s, \qquad c = \frac{T_0}{g\rho},$$

the last equation giving the value we now assign to c.

To integrate equation (3), which we now write in the form

$$(4) \qquad \frac{d^2y}{dx^2} = \frac{1}{c}\sqrt{1 + \left(\frac{dy}{dx}\right)^2},$$

we observe that (4) is a differential equation of the second order, in which the variables x and y do not both enter explicitly. On setting, therefore, by § 5

$$(5) \qquad p = \frac{dy}{dx},$$

equation (4) becomes

$$(6) \qquad \frac{dp}{dx} = \frac{\sqrt{1 + p^2}}{c}.$$

This is a differential equation of the first order, in which the variables are *separable*, § 2:

$$\frac{dx}{c} = \frac{dp}{\sqrt{1 + p^2}}.$$

* We use the notation D_x for the derivative advisedly; for, the formulation of the physical problem which culminates in equation (2) leads to *derivatives*, NOT to *differentials*. The latter are introduced later for purely *analytical* reasons. Thus the derivative expresses the thought of physics; the differential is the tool of mathematics.

Hence

$$(7)\qquad \frac{x}{c}=\int\frac{dp}{\sqrt{1+p^2}}=\log\left(p+\sqrt{1+p^2}\right)+C.$$

At A, $x=0$ and $p=0$. Hence $C=0$.

Equation (7) is equivalent to the following

$$(8)\qquad \sqrt{1+p^2}+p=e^{\frac{x}{c}}.$$

In order to solve this equation for p, we could clear of radicals and proceed as in elementary algebra; but the following method is more elegant. Take the reciprocal of each side of (8):

$$\frac{1}{\sqrt{1+p^2}+p}=e^{-\frac{x}{c}}.$$

And now rationalize the denominator by multiplying numerator and denominator by $\sqrt{1+p^2}-p$. Thus we find:

$$(9)\qquad \sqrt{1+p^2}-p=e^{-\frac{x}{c}}.$$

On subtracting (9) from (8) we have:

$$(10)\qquad 2p=e^{\frac{x}{c}}-e^{-\frac{x}{c}},\qquad \text{or}\qquad \frac{dy}{dx}=\frac{1}{2}\left(e^{\frac{x}{c}}-e^{-\frac{x}{c}}\right).$$

We can now find y:

$$y=\int\tfrac{1}{2}\left(e^{\frac{x}{c}}-e^{-\frac{x}{c}}\right)dx=\frac{c}{2}\left(e^{\frac{x}{c}}+e^{-\frac{x}{c}}\right)+K.$$

To determine K, observe that, when $x=0$, $y=c$. Hence $K=0$ and we have, as the *Equation of the Catenary:*

$$(11)\qquad y=\frac{c}{2}\left(e^{\frac{x}{c}}+e^{-\frac{x}{c}}\right).$$

All catenaries are similar. When $c=1$, we have the hyperbolic cosine, Chap. XX, § 9.

$$(12)\qquad y=\operatorname{ch}x,$$

and, generally,

$$y=c\operatorname{ch}\frac{x}{c}.$$

EXERCISES

1. *Suspension Bridge.* Find the curve in which the cables of a suspension bridge hang, when only the weight of the roadway is taken into account. *Ans.* A parabola

Note. It is assumed that the roadway is straight and horizontal, and not subject to any stress of bending; that the weight per running foot is uniform; that the vertical rods which connect the cables with the roadbed are so near together as to form approximately a continuous system; that the weight of these rods and of the cables is negligible, compared with that of the roadbed; and that the cables are perfectly flexible.

2. *Japanese Screen.* A bamboo screen for a hall closet is made by suspending slender rods of rattan, all of the same diameter, from a string, to which each rod is knotted, and letting them hang down so that each rod just touches the floor, two consecutive rods just touching each other. Find the curve in which the string hangs, if the diameter of the rods is negligible.

Ans. $y = \frac{h}{2}(e^{\frac{x}{c}} + e^{-\frac{x}{c}}), \qquad c^2 = \frac{T_0}{w}.$

3. *Suspension Bridge; General Case.* If, in Question 1, all three weights — cables, rods, and roadbed — be taken into account, find the differential equation satisfied by the cables.

Ans. $\frac{d^2y}{dx^2} = w_1 y + w_2\sqrt{1 + \frac{dy^2}{dx^2}} + w_3$

4. Show that a catenary of variable, but continuous, density is determined by the differential equation

$$\frac{d^2y}{dx^2} = \frac{g\rho}{T_0}\sqrt{1 + \frac{dy^2}{dx^2}}.$$

5. Show that a non-homogeneous heavy string can hang in equilibrium in an arc of the circle

$$x^2 + y^2 = a^2,$$

if the density is equal to $aT_0/[g(a^2 - x^2)]$.

6. Prove that the forces which act on an arc of a catenary limited by the vertex are such that the algebraic sum of their moments about the vertex is nil.

Hence show that the sum of the moments of the forces acting on an arbitrary arc is nil, no matter about what point the moments be taken.

7. Show that the centre of gravity of an arbitrary arc of a catenary lies directly above the point of intersection of the tangents drawn at the extremities.

7. Continuation; Discussion of the Catenary. The lowest point of the catenary is called the *vertex*, and the line coinciding with the axis of x, the *directrix*. The curve is symmetric in the vertical through the vertex.

Mechanical Meaning of the Directrix. If a peg, of negligible radius, be placed at any point $P:(x, y)$ of the curve, the catenary secured at P from slipping down, the string cut considerably above P and allowed to hang straight down; and if, now, this free end be cut off at the directrix; then the string, assumed smooth, can be released at P and it will not slip.

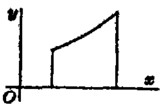


FIG. 76

For, the tension in the free end at P will be simply the weight of a segment y units long, or $g\rho y$. On the other hand, the tension at any point is, from (1),

$$T = T_0 \sec\tau = T_0 \frac{ds}{dx}.$$

Now, from (8) and (9), and (11),

$$\frac{ds}{dx} = \sqrt{1+p^2} = \tfrac{1}{2}(e^{\frac{x}{c}} + e^{-\frac{x}{c}}) = \frac{y}{c}.$$

Substituting for c its value from (3), we have:

(13) $$T = g\rho y,$$

and this completes the proof.

Thus we see that we may apply smooth pegs at any two points, hold the catenary against them, and cut the string so that each end will just reach to the directrix. On releasing the string at the pegs, it will not tend to slip.

The Arc and the Tension. The length of the arc, measured from the vertex, is

(14) $$s = \frac{c}{2}(e^{\frac{x}{c}} - e^{-\frac{x}{c}}) = c \operatorname{sh}\frac{x}{c}.$$

The tension has the value (13), or

(15) $$T = g\rho y = \frac{T_0}{2}(e^{\frac{x}{c}} + e^{-\frac{x}{c}}) = T_0 \operatorname{ch}\frac{x}{c}.$$

Problem. A chain 32 ft. long has its ends fastened at the same level to two posts 30 ft. apart. To find the dip in the chain.

This problem was studied in the *Introduction to the Calculus*, p. 174, Ex. 2. The determination of the constant c (there denoted by a) leads to an equation,

$$e^x - e^{-x} - \tfrac{32}{15}x = 0, \qquad x = \frac{15}{c},$$

which cannot be solved by the ordinary methods of algebra. That it has a solution, is clear from the fact that a continuous function which changes sign must pass through the value 0. The higher methods for solving this and similar equations, which were set forth in that chapter, lead readily to the result that $c = 23.9$, and hence the dip is found to be 4.89 ft.

EXERCISES

1. Two smooth pegs at the same level are $2a$ feet apart. Show that the shortest string which can be hung over them so as not to slip when released is of length $2l = 2ea$.

2. If the pegs in the preceding question are 2 ft. apart, and if the string is 6 ft. long, show that it can be hung over them in two, and only two, catenaries. Determine the vertex and the directrix of each.

3. Find the tension at the lowest point of the chain in the Problem of the text, if one foot of the chain weighs 4 lbs.

4. What should be the length of the chain in the Problem of the text, that the dip be precisely 1 ft.? Show that c is given by the equation

$$\operatorname{ch} x = 1 + \tfrac{1}{15}x, \qquad x = \frac{15}{c}.$$

Solve this equation for x by means of Peirce's Tables, p. 124, to two significant figures; $x = .13$.

By means of the series

$$\operatorname{ch} x = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots$$

obtain the approximation $x = \frac{2}{15}$. Then, setting $x = \frac{2}{15} - \epsilon$, obtain an approximation for ϵ, and show that $x = .1331$, to four significant figures.

5. Find the tension at the lowest point of the chain in the preceding question.

6. Show that all catenaries having the axis of x as their directrix, the vertex of each catenary being on the positive axis of y, lie above the lines $y = x\dfrac{\operatorname{ch}\lambda}{\lambda}$ and $y = -x\dfrac{\operatorname{ch}\lambda}{\lambda}$, where λ is the positive root of the equation

$$x = \coth x,$$

each catenary being tangent to each of these lines.

7. Two smooth pegs are at the same level and 2 ft. apart. Find all the positions of equilibrium for a loop (= closed string) 6 ft. long, which is hung over them.

8. **Rope round a Post.** Let it be required to find the law of tension in a rope which is wound round a post and is just on the point of slipping.

We have to do here with a flexible inextensible weightless string wound on a rough circular drum. Consider an arc PP'. Imagine this arc now to be frozen, so that we have a rigid body to deal with. Isolate this system. The forces acting on it are: (*i*) T at P; (*ii*) $T' = T + \Delta T$ at P'; and (*iii*) the reaction of the drum, which is a force S, inclined toward the direction opposite to that in which the string tends to slip. Let S make an angle λ' with the radius OP produced. *Then the physical law is that*

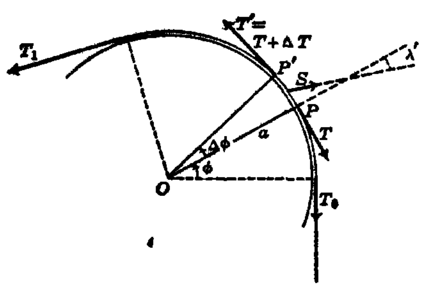


FIG. 77

(1) $$\lim_{P' \doteq P} \lambda' = \lambda,$$

where λ is the angle of friction, or $\mu = \tan\lambda$.

In order to make clear to ourselves the plausibility of this law, imagine a *heavy* chain to be laid out straight on a rough floor, and to be pulled so that it is just on the point of slipping. Consider an arc (*i.e.* a segment) PP' of this chain. The forces acting on PP' are: T, T', W, and S. And now, by the ordinary law of friction, we have, that S makes precisely the angle of friction, $\lambda = \tan^{-1}\mu$, with the normal to PP'.

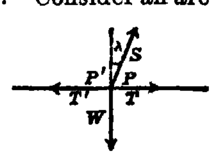


FIG. 78

The situation in the actual case is similar. Here, gravity is replaced by the pressure of the curved surface of the drum against the weightless arc PP' of the string. When the arc is short, it is like the straight line segment of the chain, and it is clear* that S will make an angle λ' with OP produced, which is nearly equal to λ.

* Not from mathematical reasoning, but from physical intuition. We are stating here the *physical hypothesis* which we lay down, and on which the whole treatment of the problem turns.

We can now proceed with the mathematical treatment of the problem. Resolving the forces along the tangent and normal at P, we have:

(a) $$T' \cos \Delta\phi = T + S \sin \lambda';$$

(b) $$T' \sin \Delta\phi = S \cos \lambda'.$$

Hence

$$\frac{T' \cos \Delta\phi - T}{T' \sin \Delta\phi} = \tan \lambda'$$

or

$$T' \cos \Delta\phi - T = T' \tan \lambda' \sin \Delta\phi.$$

Write the left-hand side of this equation in the form:

$$(T + \Delta T) \cos \Delta\phi - T = \Delta T \cos \Delta\phi - T(1 - \cos \Delta\phi).$$

We are now ready to divide through by $\Delta\phi$ and take limits:

$$\frac{\Delta T}{\Delta\phi} \cos \Delta\phi - T\,\frac{1 - \cos \Delta\phi}{\Delta\phi} = T' \tan \lambda' \frac{\sin \Delta\phi}{\Delta\phi};$$

$$\left(\lim_{\Delta\phi \doteq 0} \frac{\Delta T}{\Delta\phi}\right)\left(\lim_{\Delta\phi \doteq 0} \cos \Delta\phi\right) - T \lim_{\Delta\phi \doteq 0} \frac{1 - \cos \Delta\phi}{\Delta\phi}$$
$$= \left(\lim_{\Delta\phi \doteq 0} T'\right)\left(\lim_{\Delta\phi \doteq 0} \tan \lambda'\right)\left(\lim_{\Delta\phi \doteq 0} \frac{\sin \Delta\phi}{\Delta\phi}\right).$$

By the *Introduction to the Calculus*, Chap. V, § 3, we have:

$$\lim_{\Delta\phi \doteq 0} \frac{1 - \cos \Delta\phi}{\Delta\phi} = 0, \qquad \lim_{\Delta\phi \doteq 0} \frac{\sin \Delta\phi}{\Delta\phi} = 1;$$

and from the physical law,

$$\lim_{\Delta\phi \doteq 0} \tan \lambda' = \tan \lambda = \mu.$$

Hence, finally:

(2) $$D_\phi T = \mu T.$$

Again it is a *derivative* that expresses the *physical* thought of the problem. In order to manipulate *mathematically* the result, we introduce *differentials*:

(3) $$\frac{dT}{d\phi} = \mu T.$$

This differential equation can be integrated by separating the variables, § 2:

$$\frac{dT}{T} = \mu\, d\phi;$$

$$\int \frac{dT}{T} = \int \mu d\phi;$$

$$\log T = \mu\phi + C.$$

When $\phi = 0$, $T = T_0$. Hence $C = \log T_0$, and we have as the law of tension:

(4) $$T = T_0 e^{\mu\phi}.$$

It will be observed that the result is independent of the radius of the drum.

EXERCISES

1. The *Miles Standish* is docking at Nantasket and a longshoreman is holding her by a hawser wound round a post of the wharf. If the coefficient of friction is $\frac{1}{3}$, and if the steamer is pulling with a force of 5 tons, how many turns of the hauser are necessary, in order that the man need pull with a force of only 50 lbs.?

Ans. Slightly over two and a half.

2. If the coefficient of friction of a band brake is $\frac{1}{2}$, show that the brake will be nearly six times as effective when applied to a complete circumference, as when applied only to half a circumference.

9. Heavy Strings on Surfaces, Rough or Smooth. Let a heavy chain of continuous density, ρ, rest on a smooth surface and lie in a vertical plane. The forces acting on an arc PP', which we isolate and assume to become rigid, are: T, $T' = T + \Delta T$, S, and $g\bar{\rho}\Delta s$, where $\bar{\rho}$ denotes the mean, or average, density of the arc. Let the angle from S to the normal at P be ϵ. *Then the physical law is that*

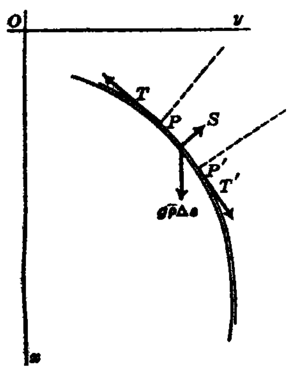


FIG. 79

(1) $$\lim_{P' \doteq P} \epsilon = 0.$$

In fact, it is clear* that the direction of S must lie between the two extreme normals — the one at P and the one at P'.

Resolving along the tangent and normal at P, we find:

(a) $$T' \cos \Delta\tau + g\bar{\rho}\Delta s \cos \tau + S \sin \epsilon = T;$$

(b) $$T' \sin(-\Delta\tau) + g\bar{\rho}\Delta s \sin \tau = S \cos \epsilon.$$

* This is, of course, only another form of statement for the physical law.

On eliminating S between (a) and (b), we have:

$$\frac{T' \cos \Delta\tau - T + g\bar{\rho}\Delta s \cos\tau}{T' \sin \Delta\tau - g\bar{\rho}\Delta s \sin\tau} = \tan\epsilon,$$

or

$$\frac{\Delta T}{\Delta x}\cos\Delta\tau - T\,\frac{1-\cos\Delta\tau}{\Delta\tau}\frac{\Delta\tau}{\Delta x} + g\bar{\rho}\frac{\Delta s}{\Delta x}\cos\tau$$

$$= \tan\epsilon\left(T'\frac{\sin\Delta\tau}{\Delta\tau}\frac{\Delta\tau}{\Delta s}\frac{\Delta s}{\Delta x} - g\bar{\rho}\frac{\Delta s}{\Delta x}\sin\tau\right).$$

Now allow Δx to approach 0. The right-hand side of the equation approaches 0, for $\lim\epsilon = 0$ and the parenthesis approaches a limit. Hence the limit of the left-hand side is 0, also. This latter limit has the form below, or:*

(2) $$D_x T + g\rho D_x s \cos\tau = 0,$$

or

(3) $$\frac{dT}{dx} = -g\rho.$$

If $\rho = \text{const.}$, (3) gives:

(4) $$T_0 - T = g\rho(x - x_0).$$

Suppose the lower end of the chain has the abscissa x_1. The tension there is $T = 0$, and so the tension T_0 at the upper end is

$$T_0 = g\rho(x_1 - x_0).$$

This is precisely the value that the tension would have if a length $x_1 - x_0$ of the chain hung vertically downward. We see, then, that a chain will rest in equilibrium on a smooth surface if it is allowed to hang over the upper part of the surface, the free end reaching down to the level of the end which is on the surface.

We have tacitly assumed that the curve is always concave toward the negative axis of y, and that $0 < \tau < \pi/2$. The treatment holds, however, with slight modification for a heavy string in a smooth tube.

If ρ is very small and $x - x_0$ only moderately large, T is always nearly equal to T_0. We see, therefore, that it must be a physical

* It might seem as if this equation could have been inferred at once from (a) alone, written in the form:

$$\Delta T \cos\Delta\tau - T(1 - \cos\Delta\tau) + g\bar{\rho}\,\Delta s \cos\tau + S\sin\epsilon = 0,$$

since $\lim \sin\epsilon = 0$. But it is not clear from (a) alone that $S/\Delta x$ approaches a limit (the essential thing is that this variable remain finite), when Δx approaches 0.

fact that the tension in a weightless string which passes over a smooth surface is the same throughout. This checks with common sense. The tension in a rope which passes through a block is the same on both sides of the pulley, — or would be so if the pulley were frictionless and the rope perfectly flexible and weightless.

EXERCISES

1. On a rough circular cylinder with horizontal axis is placed a chain, one end of which is at the level of the axis; the other, hanging down to a distance l below the axis, and the chain is just on the point of slipping. Show that the tension, T, satisfies the differential equation:

$$\frac{dT}{d\phi} - \mu T = g\rho a(\cos\phi + \mu\sin\phi).$$

Assume the chain to be a uniform, flexible, inextensible string, lying in a plane perpendicular to the axis of the cylinder.

2. Prove that, in Question 1,

$$l = \frac{2\mu a}{1+\mu^2}(1 + e^{\pi\mu}).$$

3. A piece of the chain of Question 1, equal in length to a quadrant of the cross-section of the cylinder, is laid along such a quadrant, the lower end of the chain being at the level of the axis of the cylinder. Show that the least value which μ may have, if the chain is not to slip, is given by the equation:

$$\tan 2\lambda = e^{\frac{\pi}{2}\tan\lambda}.$$

4. Solve this equation for μ to four significant figures.

Ans. $\mu = 0.7322$.

5. Two smooth circular cylinders, external to each other, have their axes horizontal. A heavy chain is hung over them, and is in equilibrium in a plane perpendicular to their axes. Show that its ends lie in the directrix of the catenary in which the part of the chain between the cylinders hangs.

6. Show that, in the problem studied in the text, $S/\Delta s$ approaches a limit when P' approaches P. This limit, σ, may be thought of as the *specific reaction*, or the pressure per unit of length, which the drum exerts on the string.

7. It was assumed in deducing equation (3) that $\Delta T/\Delta x$ approaches a limit, and then that limit was computed. From (a) and

(*b*) prove (*i*) that ΔT approaches 0 when Δx approaches 0; and (*ii*) that $\Delta T/\Delta x$ approaches a limit.

8. A heavy string of continuous density is in equilibrium in a smooth tube in the form of a twisted curve. Show that the tension, T, satisfies the differential equation (3), the axis of x being vertical and positive downward. If $\rho =$ const., equation (4) holds in this case, too.

9. *Pile of Theme Paper.* Suppose that a pile of theme paper is pierced by a hole near the middle of an edge; that a string is inserted; and that the paper is hung up, being prevented from bulging by two vertical walls. If the string, paper, and walls are all smooth, find the curve in which the string hangs.

Ans. An arc of a circle.

Note. Although perfect smoothness cannot be attained physically, still, a close approximation to the conditions of the problem can be realized by hanging the paper up in a freight car. The jarring will cause the paper to adjust itself as prescribed. Thin metal plates, all of the same weight, would be better adapted to the purpose.

10. What will be the differential equation of the curve of the string in the preceding question if the paper is trimmed so that the lower edges all lie at the same level?

Ans. $$\frac{d}{dx}\frac{\dfrac{dy}{dx}}{\sqrt{1+\left(\dfrac{dy}{dx}\right)^2}}=\frac{y}{c^2}, \qquad c^2=\frac{T_0}{w}.$$

11. *The Hydrostatic Arch.* Consider a canal for carrying water across a ravine, the bottom of the canal being an arch. What must be the shape of the latter, in order that it may not tend to bend at any point? Neglect the weight of the arch in comparison with the weight of the water it supports.

Let the axis of x be chosen in the surface of the water, at right angles to the direction of the canal, and let the axis of y be positive downward, passing through the top of the arch. Then

$$\frac{d}{dx}\frac{p}{\sqrt{1+p^2}}=\frac{y}{c^2}, \qquad p=\frac{dy}{dx}.$$

12. Integrate the differential equation of Question 11.

Ans. $x=\displaystyle\int\frac{2c^2+h^2-y^2}{\sqrt{(y^2-h^2)(4c^2+h^2-y^2)}}dy$, where h denotes the minimum depth of the water.

If $v_D > v_M$, then $c < 1$, and the duck will make port. If, however, $v_D < v_M$, the duck will be carried down stream and approach the opposite bank asymptotically. Finally, if $v_D = v_M$, then $c = 1$, and the path is a parabola. The duck will sidle up toward a point half as far down the bank as the breadth of the stream — much as the Rhine steamers make a landing.

(c) *The Dog and His Master.* A dog, out in a field, sees his master walking along the road and runs toward him. Find the path of the dog. It is assumed that the dog always heads straight for his master, that each moves at a uniform rate, and that the road is straight.

The same figure can be used as in the case of the tractrix, but with a different interpretation. For here,

$$s = v_D t, \qquad \overline{OQ} = v_M t,$$

and it is expedient to observe that

$$x = \overline{OQ} - \overline{MQ} = v_M t + y \cot \tau.$$

Thus

$$x = cs + y\frac{dx}{dy}, \qquad c = \frac{v_M}{v_D}.$$

Differentiating with respect to y and observing that ds/dy is negative, we find:

$$(5) \qquad y\frac{d^2x}{dy^2} = c\sqrt{1 + \left(\frac{dx}{dy}\right)^2}.$$

The integral of this equation has been found for the case that $c \neq 1$ and $y = 1$, $dx/dy = 0$, when $x = 0$; § 5, Ex. 2:

$$(6) \qquad x = \frac{1}{2}\left\{\frac{y^{1+c}}{1+c} - \frac{y^{1-c}}{1-c} + \frac{2c}{1-c^2}\right\},$$

and since, for a given value of c, all curves are similar, this one gives the shape for the whole class.

If $c < 1$, the dog overtakes his master at the point of the road for which $x = c/(1 - c^2)$. If $c > 1$, the dog approaches the road asymptotically. The case $c = 1$ presents no difficulty; but equation (6) is replaced by one in which a logarithmic term appears.

A number of further problems similar in character to those discussed here are given in Tait & Steele's *Dynamics of a Particle*, Chap. I.

EXERCISES

1. A man swims across a river, always heading straight for the opposite bank. If the current is such that he is carried down stream with a velocity proportional to his distance from the nearer bank, find his path. *Ans.* A curve made up of two equal parabolic arcs.

2. A circular turn-table rotates about its axis with uniform velocity. An ant steps on at the outer edge and crawls straight toward a light at the centre of the table. Find the path of the ant in space.

Ans. $r = a(1 - c\theta)$.

3. If the sun is setting in the west and the ant boards the turn-table at its most easterly point and then always crawls straight toward the sun, show that the ant will describe an arc of a circle.

4. If in Question 2 the light had been at a point fixed in space, on the circumference of the turn-table and diametrically opposite the point at which the ant steps on, obtain the differential equation of the path of the ant.

II. Linear Equations of the Second Order, and Higher

11. Elementary Theorems. The differential equation

$$\frac{d^n y}{dx^n} + P_1 \frac{d^{n-1}y}{dx^{n-1}} + \cdots + P_{n-1}\frac{dy}{dx} + P_n y = R, \tag{1}$$

in which the coefficients $P_1, \cdots, P_n, R$, are given functions of x, which do not depend on y, is called a *linear differential equation*, because it is linear in y and its derivatives. If $R \equiv 0$:

$$\frac{d^n y}{dx^n} + P_1 \frac{d^{n-1}y}{dx^{n-1}} + \cdots + P_n y = 0, \tag{2}$$

the equation is said to be *homogeneous;* otherwise, *non-homogeneous.* The homogeneous equations form far and away the more important class.

Theorem I. *If y_1 be a solution of the homogeneous linear differential equation* (2), *then cy_1, where c is any constant, is also a solution.*

By hypothesis, y_1 satisfies equation (2), *i.e.*

$$\frac{d^n y_1}{dx^n} + P_1 \frac{d^{n-1}y_1}{dx^{n-1}} + \cdots + P_n y_1 = 0 \tag{3}$$

is a true equation. We wish to prove that cy_1 also satisfies equation (2), *i.e.* that

(4) $$\frac{d^n(cy_1)}{dx^n} + P_1\frac{d^{n-1}(cy_1)}{dx^{n-1}} + \cdots + P_n(cy_1) = 0$$

is a true equation. It is clear how to draw the inference.

THEOREM II. *If y_1 and y_2 be two solutions of the homogeneous linear differential equation* (2), *then their sum, $y_1 + y_2$, is also a solution.*

The proof is similar to that of Theorem I, and is left to the student.

Linearly Independent Functions. If n functions, $f_1(x), \cdots, f_n(x)$, are connected by an identical relation of the form:

(5) $$c_1f_1(x) + \cdots + c_nf_n(x) = 0,$$

where the c's are constants not all 0, the functions are said to be *linearly dependent.* If no such relation between the functions exists, they are called *linearly independent.*

Thus, for $n = 3$, the functions

$$f_1(x) = \sin x, \qquad f_2(x) = \cos x, \qquad f_3(x) = \sin(x + \alpha)$$

are linearly dependent; for

$$f_1(x)\cos\alpha + f_2(x)\sin\alpha - f_3(x) \equiv 0.$$

For an arbitrary value of n the first n powers of x, namely, $x^0 = 1$, $x^1 = x$, $x^2, \cdots, x^{n-1}$, form a set of n linearly independent functions; for the function

$$c_0 + c_1x + c_2x^2 + \cdots + c_{n-1}x^{n-1}$$

vanishes identically when and only when each coefficient is 0.

Existence Theorem. It is shown in the theory of linear differential equations that, if the coefficients of the homogeneous linear differential equation (2) be continuous in an interval $a \leqq x \leqq b$, there exist n linearly independent solutions, $y_1, \cdots, y_n$, each defined throughout the interval.

From Theorems I and II it appears that the function

(6) $$y = c_1y_1 + c_2y_2 + \cdots + c_ny_n$$

is also a solution, where the c's are any constants.

And now it is shown, furthermore, in the theory of differential equations that, conversely, every solution of (2) in the above interval can be written in the form (6).

The Non-Homogeneous Equation (1). The solution of this equation can be referred to that of the corresponding homogeneous equation (2) whenever one single particular solution of (1) can be found, as is stated in precise form by the following

THEOREM. *Let Y be a particular solution of* (1); *i.e. a function which satisfies* (1), *but contains no constants of integration. Then the general solution of* (1) *is*

$$y = Y + c_1y_1 + \cdots + c_ny_n,$$

where $y_1, \cdots, y_n$ are n linearly independent solutions of the corresponding homogeneous differential equation (2), *and $c_1, \cdots, c_n$ are arbitrary constants.*

By hypothesis, we have the equation

$$\frac{d^nY}{dx^n} + P_1\frac{d^{n-1}Y}{dx^{n-1}} + \cdots + P_nY = R. \tag{7}$$

Now let y be any solution whatever of (1). Subtract (7) from (1); then

$$\frac{d^n(y-Y)}{dx^n} + P_1\frac{d^{n-1}(y-Y)}{dx^{n-1}} + \cdots + P_n(y-Y) = 0,$$

i.e. the function $y - Y$ satisfies (2). It can, therefore, be written in the form (6):

$$y - Y = c_1y_1 + c_2y_2 + \cdots + c_ny_n,$$

and the theorem is proved.

EXERCISES

Show that the following functions are linearly independent when $n = 2$:

1. $\sin x, \quad \cos x.$
2. $e^{-kt}\sin pt, \quad e^{-kt}\cos pt.$
3.* $x, \quad e^x.$
4. $e^x, \quad \sin x.$
5. $e^{px}, \quad e^{qx}, \quad p \neq q.$
6. $e^{mx}, \quad xe^{mx}.$

Are the following functions linearly independent?

7. $n = 3: \quad e^{\alpha x}, \quad e^{\beta x}\cos px, \quad e^{\beta x}\sin px.$
8. $n = 4: \quad \sin px, \quad \cos px, \quad \sin qx, \quad \cos qx.$
9. $n = 4: \quad e^{\alpha x}\sin px, \quad e^{\alpha x}\cos px, \quad e^{\beta x}\sin qx, \quad e^{\beta x}\cos qx.$

10. By a *simultaneous system of n linear differential equations of the first order* is meant:

$$\frac{dy_k}{dx} = \sum_{i=1}^{n} a_{ik}y_i + a_k, \qquad k = 1, \cdots, n.$$

* *Suggestion.* Assume the theorem false. Then $Ax + Be^x = 0$ for all values of x; and now differentiate.

The coefficients, a_{ik} and a_k, are any continuous functions of x. The system is said to be *homogeneous* if $a_k = 0$, $k = 1, \cdots, n$.

Write out such a system (both non-homogeneous and homogeneous) for $n = 2$ and $n = 3$.

By a *solution* of such a system is meant a set of n functions.

$$y_k = f_k(x), \qquad k = 1, \cdots, n,$$

which satisfy the given system.

State and prove Theorems I and II for a homogeneous system, and the last Theorem of the text for a non-homogeneous system.

11. Show that the linear differential equation

$$x^2 \frac{d^2 y}{dx^2} + Px\frac{dy}{dx} + Qy = R$$

goes over by the substitution $x = e^t$ into the linear differential equation

$$\frac{d^2 y}{dt^2} + (P-1)\frac{dy}{dt} + Qy = R.$$

Extend the theorem to linear differential equations of the n-th order.

12. Constant Coefficients. We begin with the case of the homogeneous differential equation of the second order,

$$\text{(1)} \qquad \frac{d^2 y}{dx^2} + 2\alpha\frac{dy}{dx} + \beta y = 0,$$

where α, β are given constants. It was early observed that the function e^{mx} is a solution of this differential equation if m is a root of the quadratic equation

$$\text{(2)} \qquad m^2 + 2\alpha m + \beta = 0.$$

For, compute the left-hand side of (1) when $y = e^{mx}$. Here,

$$\frac{dy}{dx} = me^{mx}, \qquad \frac{d^2 y}{dx^2} = m^2 e^{mx},$$

and hence

$$\frac{d^2 y}{dx^2} + 2\alpha\frac{dy}{dx} + \beta y = e^{mx}(m^2 + 2\alpha m + \beta).$$

Thus we see, for example, that the differential equation

$$\text{(3)} \qquad \frac{d^2 y}{dx^2} - 5\frac{dy}{dx} + 6y = 0$$

leads to the quadratic

$$m^2 - 5m + 6 = 0,$$

whose roots are $m_1 = 2$, $m_2 = 3$. Hence two solutions of (3) are e^{2x} and e^{3x}. These are evidently linearly independent, and so the general solution of (3) is

$$y = Ae^{2x} + Be^{3x}.$$

Imaginary Roots of (2). Suppose, however, that (2) has no real roots, *i.e.* suppose that, on writing down the formal solution of (2),

$$m = -\alpha \pm \sqrt{\alpha^2 - \beta},$$

it turns out that $\alpha^2 - \beta < 0$. The roots of (2) are then *imaginary*, as the mathematicians of the eighteenth century said. They can be expressed in the form

$$(4) \qquad m_1 = -\alpha + \gamma\sqrt{-1}, \qquad m_2 = -\alpha - \gamma\sqrt{-1},$$

where $\gamma = \sqrt{\beta - \alpha^2}$. The mathematicians of that time did not hesitate to work with imaginary expressions like the above, even though they had no clear idea of what they mean, *i.e.* how to define them. They reasoned as follows. Since $e^{u+v} = e^u e^v$ when u and v are real, the expression

$$e^{mx} = e^{-\alpha x - \gamma x\sqrt{-1}}$$

must be the same thing as the product

$$e^{-\alpha x} e^{-\gamma x\sqrt{-1}},$$

and so the question reduces itself to that of finding out what $e^{\phi\sqrt{-1}}$ means, where ϕ is a real number.

Now, the mathematicians of that time were very well acquainted with the expansions of the functions e^x, $\sin x$, $\cos x$ by Taylor's Theorem:

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots.$$

What could be more natural, therefore, than to ask the series what $e^{\phi\sqrt{-1}}$ means? On setting $x = \phi\sqrt{-1}$ in the above development of e^x and reducing the result by means of the relations

$$(\sqrt{-1})^1 = \sqrt{-1}, \qquad (\sqrt{-1})^2 = -1, \qquad (\sqrt{-1})^3 = -\sqrt{-1},$$

$$(\sqrt{-1})^4 = 1, \qquad (\sqrt{-1})^{4k+l} = (\sqrt{-1})^l,$$

$$k = 1, 2, \cdots; \qquad l = 1, 2, 3, 4,$$

it appears that

$$e^{\phi\sqrt{-1}} = 1 - \frac{\phi^2}{2!} + \frac{\phi^4}{4!} - \cdots$$

$$+ \sqrt{-1}\left(\phi - \frac{\phi^3}{3!} + \frac{\phi^5}{5!} - \cdots\right).$$

But these are precisely the series for $\cos\phi$ and $\sin\phi$. And so the formal work indicates that we ought to consider $e^{\phi\sqrt{-1}}$ as equivalent to $\cos\phi + \sqrt{-1}\sin\phi$:

(5) $$e^{\phi\sqrt{-1}} = \cos\phi + \sqrt{-1}\sin\phi.$$

We cannot insist too strongly at this point that we have not *proved* the equation (5). How can we prove anything about imaginaries before we have a definition for them? We can, however, act as if we had proved equation (5) and go on and reduce the expression for e^{m_1x} accordingly. We had found that this is to be considered as the product $e^{-ax}e^{\gamma x\sqrt{-1}}$, and hence we interpret e^{m_1x} in the light of (5) by the equation:

$$e^{m_1x} = e^{-ax}\cos\gamma x + \sqrt{-1}\,e^{-ax}\sin\gamma x.$$

Similarly, $$e^{m_2x} = e^{-ax}\cos\gamma x - \sqrt{-1}\,e^{-ax}\sin\gamma x.$$

Now, the sum of two solutions of the differential equation (1) is, by Theorem II of § 11, a solution. And the sum of e^{m_1x} and e^{m_2x}, if it means anything, means the function

$$2e^{-ax}\cos\gamma x.$$

But this is a real function, and it may be a solution of (1), in spite of the doubtful character of its pedigree. Try it and see. Dropping the factor 2, set

(6) $$y = e^{-ax}\cos\gamma x;$$

then $$\frac{dy}{dx} = -\alpha e^{-ax}\cos\gamma x - \gamma e^{-ax}\sin\gamma x,$$

$$\frac{d^2y}{dx^2} = (\alpha^2 - \gamma^2)e^{-ax}\cos\gamma x + 2\alpha\gamma e^{-ax}\sin\gamma x,$$

and on adding these equations after multiplying the first by β, the second by 2α, and observing that $\gamma^2 = \beta - \alpha^2$, we find that

$$\frac{d^2y}{dx^2} + 2\alpha\frac{dy}{dx} + \beta y \equiv 0,$$

i.e. the function (6) is *proved* by direct substitution, to be a solution of (1).

In a similar manner we can find a second solution of (1). Subtract e^{m_2x} from e^{m_1x}; the result is

$$2\sqrt{-1}\,e^{-\alpha x}\sin\gamma x.$$

It is true that this result is imaginary, but only through the presence of an imaginary constant factor, $2\sqrt{-1}$. Suppress this factor and consider the function

$$y = e^{-\alpha x}\sin\gamma x. \tag{7}$$

On substituting this function into the given differential equation, as was done with the function (6), we find that (7) also satisfies that equation, and thus we have the best of all proofs that (7) is a solution — that of direct substitution. For, a function that satisfies a differential equation is a solution, no matter how obscure its origin; and one that does not satisfy it is not a solution, no matter how illustrious its pedigree may seem to have been.

We have introduced this bit of eighteenth century mathematics partly to give a *motif* for the two solutions (6) and (7); partly to show how mathematicians obtained true results from working with $\sqrt{-1}$, long before they knew how to define that number. They divined its importance, but they did not yet have the vision to give it existence through definition, as is seen from a remark of Leibniz in the year 1702*: "Die imaginären Zahlen sind eine feine und wunderbare Zuflucht des göttlichen Geistes, beinahe ein Amphibium zwischen Sein und Nichtsein."

Equations of the n-th Order. The method can be extended at once to the equation

$$\frac{d^ny}{dx^n} + \alpha_1\frac{d^{n-1}y}{dx^{n-1}} + \cdots + \alpha_n y = 0, \tag{8}$$

where the α's are real constants. On substituting $y = e^{mx}$ we find that this function is a solution provided m is a root of the algebraic equation

$$m^n + \alpha_1 m^{n-1} + \cdots + \alpha_n = 0. \tag{9}$$

If this equation has n real and distinct roots, $m_1, \cdots, m_n$, the general solution of (8) will be

$$y = c_1e^{m_1x} + \cdots + c_ne^{m_nx}. \tag{10}$$

If one of the roots of (9) is imaginary,

$$m_1 = p + q\sqrt{-1},$$

* Klein, *Elementarmathematik vom höheren Standpunkte aus*, 3d ed., vol. I, p. 61.

then a second root will be the conjugate imaginary,

$$m_2 = p - q\sqrt{-1}.$$

Corresponding to these roots we shall have two real solutions,

(11) $$e^{px}\cos qx, \qquad e^{px}\sin qx.$$

The case of equal roots of (2) or (9) will be treated in the next paragraph.

Example. Solve the differential equation

$$\frac{d^3y}{dx^3} + \frac{dy}{dx} = 0.$$

Here, equation (9) becomes

$$m^3 + m = 0.$$

The roots of this equation are

$$m_1 = 0, \qquad m_2 = \sqrt{-1}, \qquad m_3 = -\sqrt{-1}.$$

Hence $e^{m_1x} = 1$ is one solution, and two further solutions are

$$e^{px}\cos qx = \cos x, \qquad e^{px}\sin qx = \sin x.$$

The general solution is

$$y = A + B\cos x + C\sin x.$$

EXERCISES

Solve the following differential equations.

1. $\frac{d^2y}{dx^2} - y = 0.$ *Ans.* $y = Ae^x + Be^{-x}.$

2. $\frac{d^2y}{dx^2} + y = 0.$ *Ans.* $y = A\cos x + B\sin x.$

3. $\frac{d^2y}{dx^2} + 3\frac{dy}{dx} - 10y = 0.$

4. $\frac{d^2y}{dx^2} + \frac{dy}{dx} - y = 0.$

5. $\frac{d^2y}{dx^2} + 4\frac{dy}{dx} + 13y = 0.$

6. $\frac{d^2y}{dx^2} + \frac{dy}{dx} + y = 0.$

7. $\frac{d^3y}{dx^3} + y = 0.$

8. $\frac{d^3y}{dx^3} - \frac{d^2y}{dx^2} + \frac{dy}{dx} - y = 0.$

9. $\frac{d^4y}{dx^4} - y = 0,$

10. $\frac{d^4y}{dx^4} + 13\frac{d^2y}{dx^2} + 36y = 0.$

11. $\frac{d^4y}{dx^4} + \frac{dy}{dx} = 0.$

12. $\frac{d^4y}{dx^4} - 13\frac{d^2y}{dx^2} + 36y = 0,$

13. $\frac{d^6y}{dx^6} - y = 0.$

14. $\frac{d^7y}{dx^7} - \frac{dy}{dx} = 0.$

15. Show that one solution of the differential equation

$$\frac{d^n y}{dx^n} + \alpha_1 \frac{d^{n-1} y}{dx^{n-1}} + \cdots + \alpha_n y = \epsilon,$$

where $\alpha_1, \cdots, \alpha_n, \epsilon$ are constants, and $\alpha_n \neq 0$, is the function $y = \epsilon/\alpha_n$.

16. Obtain by inspection one solution of the differential equation

$$\frac{d^2 y}{dx^2} + 2\frac{dy}{dx} = 4,$$

and hence solve the equation completely.

17. In the case of a simultaneous system of homogeneous linear differential equations with constant coefficients, as

$$\frac{dy}{dx} = Ay + Bz, \qquad \frac{dz}{dx} = Cy + Dz,$$

it is reasonable to try for a solution of the form:

$$y = \lambda e^{mx}, \qquad z = \mu e^{mx}.$$

Show that two such solutions can be found if the equation

$$\begin{vmatrix} A - m & B \\ C & D - m \end{vmatrix} = 0$$

has two distinct real roots, and determine the ratio, λ/μ.

Apply your results to the case:

$$\frac{dy}{dx} = 6y - 4z, \qquad \frac{dz}{dx} = 3y - z.$$

Ans. The complete solution is:

$$y = C_1 e^{2x} + 4C_2 e^{3x}, \qquad z = C_1 e^{2x} + 3C_2 e^{3x}.$$

18. Develop the theory for the case that the quadratic in m, Question 17, has imaginary roots.

19. Extend Questions 17, 18 to the case of a system of three equations,

$$\frac{du}{dx} = A_1 u + B_1 v + C_1 w,$$

$$\frac{dv}{dx} = A_2 u + B_2 v + C_2 w,$$

$$\frac{dw}{dx} = A_3 u + B_3 v + C_3 w.$$

Hence generalize to the case of n equations.

20. Solve the differential equation (cf. Ex. 11, § 11):

$$x^2\frac{d^2y}{dx^2}+8x\frac{dy}{dx}+12y=0.$$

Ans. $y=C_1x^{-3}+C_2x^{-4}$.

13. Continuation. Equal Roots.* If equation (2) of § 12 has equal roots, then each is equal to $-\alpha$, and $y_1=e^{-\alpha x}$ is a solution of the given differential equation.

We can guess a second solution by considering the case that the roots are not quite equal, one being $-\alpha$ and the other $-\alpha+h$. Let

$$\bar{y}=e^{-\alpha x+hx}.$$

Then
$$\frac{\bar{y}-y_1}{h}=e^{-\alpha x}\frac{e^{hx}-1}{h}$$

is a second solution of the near-equation, no matter how small h be taken. Now, this function is nearly equal to $xe^{-\alpha x}$ when h is small. So we are led to try this function, and it turns out on substituting it that it does satisfy the given differential equation.

Thus we find as the general solution

$$y=e^{-\alpha x}(A+Bx).$$

If $n>2$, equation (9) of § 12 may have more than two roots equal. It is not hard now to guess by analogy what the solutions will be in this case. If m be an l-fold real root, then

$$y_1=e^{mx},\qquad y_2=xe^{mx},\quad\cdot\quad\cdot\quad\cdot,\quad y_l=x^{l-1}e^{mx}$$

will be l linearly independent solutions. If, on the other hand, m is imaginary: $m=p+q\sqrt{-1}$, then $p-q\sqrt{-1}$ will also be a root, and we have each of these roots counting l times. The functions

$$y_{2k+1}=x^ke^{px}\cos qx,\qquad y_{2k+2}=x^ke^{px}\sin qx,\qquad k=0,1,\cdots,l-1,$$

are here $2l$ linearly independent solutions.

The case that the m-equation for a simultaneous system of the type of § 12, Exs. 17 and 19, has equal roots is more complex; cf. Goursat, *Cours d'analyse*, Vol. II, 2d ed. (1911), Chap. XX, § 420, p. 483.

* This case is unimportant in practice; and yet it is necessary to treat it if the theory is to be complete. The student may safely postpone this paragraph till he needs to use it.

EXERCISES

1. Find one solution of the differential equation

$$\frac{d^3y}{dx^3} - 3\frac{d^2y}{dx^2} + 3\frac{dy}{dx} - y = 0$$

by the method of § 12, and prove by direct substitution that xe^x and x^2e^x are also solutions. What is the general solution?

Ans. $y = (c_0 + c_1x + c_2x^2)e^x$.

2. Find two solutions of the differential equation

$$\frac{d^4y}{dx^4} + 2\frac{d^2y}{dx^2} + y = 0$$

by the method of § 12, and prove by direct substitution that $x \sin x$ and $x \cos x$ are also solutions. What is the general solution?

Ans. $y = (a + bx)\cos x + (c + dx)\sin x$.

Solve completely the following differential equations.

3. $\dfrac{d^4y}{dx^4} + \dfrac{d^3y}{dx^3} = 0.$

4. $\dfrac{d^4y}{dx^4} + \dfrac{d^2y}{dx^2} = 0.$

5. $\dfrac{d^3y}{dx^3} + \dfrac{d^2y}{dx^2} - \dfrac{dy}{dx} - y = 0.$

6. $\dfrac{d^6y}{dx^6} - 2\dfrac{d^3y}{dx^3} + y = 0.$

14. Small Oscillations of a System with n Degrees of Freedom. We treat here only that part of the problem which relates to the integration of the differential equations involved.* Let the kinetic energy, T, and the force-function, U, be given by the equations:

$$T = \sum a_{ij}q_i'q_j' + T_1, \qquad U = -\sum b_{ij}q_iq_j + U_1, \qquad \begin{cases} a_{ij} = a_{ji} \\ b_{ij} = b_{ji} \end{cases}$$

The two quadratic forms are both definite, and a_{ij}, b_{ij} are constants. Lagrange's Equations:

$$\frac{d}{dt}\frac{\partial T}{\partial q_k'} - \frac{\partial T}{\partial q_k} = \frac{\partial U}{\partial q_k}$$

give:

(1) $$a_{k1}q_1'' + \cdots + a_{kn}q_n'' = -(b_{k1}q_1 + \cdots + b_{kn}q_n).$$

By means of a suitable linear transformation,

(2) $$q_k = \mu_{k1}p_1 + \cdots + \mu_{kn}p_n, \qquad k = 1, \cdots, n,$$

the two quadratic forms can be reduced to the normal form: †

* Cf. Appell, *Mécanique rationelle*, vol. 2, p. 343.

† Bôcher, *Algebra*, Chap. 13.

$$p_1'^2 + \cdots + p_n'^2, \qquad -r_1^2p_1^2 - \cdots - r_n^2p_n^2, \qquad 0 < r_k$$

Equations (1) now take on the form:

$$(3) \qquad p_k'' = -r_k^2 p_k.$$

The general solution of the system of equations (3) is obvious, namely:

$$(4) \qquad p_k = C_k \cos(r_k t + \rho_k), \qquad k = 1, \cdots, n.$$

Hence the general solution of (1) will be:

$$(5) \quad q_k = C_1\mu_{k1}\cos(r_1 t + \rho_1) + \cdots + C_n\mu_{kn}\cos(r_n t + \rho_n), \quad k = 1, \cdots, n,$$

where $C_1, \cdots, C_n$ and $\rho_1, \cdots, \rho_n$ are the $2n$ constants of integration.

From the result just obtained it is now clear how to solve equations (1) without the intervention of the transformation (2). For, equations (5) say that the general solution of (1) is put together linearly out of n solutions, each of the form:

$$(6) \qquad q_1 = \lambda_1 \cos(rt + \rho), \quad \cdots, \quad q_n = \lambda_n \cos(rt + \rho).$$

A necessary condition that (6) be a solution is that the n equations

$$(7) \qquad (b_{k1} - r^2 a_{k1})\lambda_1 + \cdots + (b_{kn} - r^2 a_{kn})\lambda_n = 0, \qquad k = 1, \cdots, n,$$

admit a solution in which the λ_i's are not all 0. Hence the determinant of these equations must vanish:

$$(8) \qquad \begin{vmatrix} b_{11} - r^2 a_{11}, & \cdots, & b_{1n} - r^2 a_{1n} \\ \cdot & \cdots & \cdot \\ b_{n1} - r^2 a_{n1}, & \cdots, & b_{nn} - r^2 a_{nn} \end{vmatrix} = 0.$$

If the r_k's are all distinct, they form precisely the n positive roots of (8) and thus (8) is seen to be a sufficient, as well as a necessary, condition for the r_k.

It is not difficult to show by a limiting process that, when two or more of the r_k in (3) are equal, these appear as multiple roots of (8), so that, in all cases, the n positive roots of (8) yield the n quantities r_k.

When r_k is a simple root of (8), equations (7), written for $r = r_k$, determine the ratios of the λ_j's uniquely, and thus, in building up the general solution (5) out of such particular solutions (6), the factor of proportionality can be merged with the coefficient C_k.

If $r = r_k$ is a multiple root of (8), of order m, then m of the λ_j's in (7) can be chosen arbitrarily. For example, any one of these m

λ_j's can be set $=1$, and the remaining $m-1$ set $=0$. We are thus led to m linearly independent solutions of (1).*

We see, therefore, how in all cases to derive from (8) and (7) n solutions of the form (6), out of which an arbitrary solution of (1) can be constructed by means of (5).

The variables p_k are known as the *normal coordinates* of the system. They are uniquely determined, save as to their order, when the r_k are all distinct; but when some of the r_k's are equal, an infinite number of different choices is possible.

III. Geometrical Interpretation. Singular Solutions

15. Meaning of a Differential Equation. Just as, in Integration, our first object was to discover the devices by which the integrals we meet in practice can be evaluated in terms of the elementary functions, so here we have studied in this chapter analogous devices for solving differential equations such as occur in physics and geometry by means of explicit formulas in the elementary functions. We came, however, to see that an integral can be considered from a higher point of view and that the integral of any continuous function always exists, regardless of whether it can be evaluated as above; namely, the *area under the curve* yields precisely the integral. Moreover, this area may in any case be approximated to by Simpson's Rule, *Introduction*, p. 344.

In the case of the differential equation

$$\frac{dy}{dx}=f(x, y) \tag{1}$$

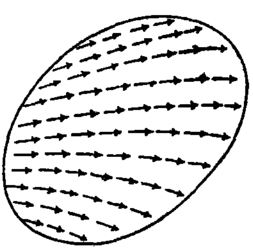


Fig. 82

the situation is similar. Suppose $f(x, y)$ to be continuous throughout a certain region S of the (x, y)-plane. Then the equation (1) assigns to each point (x, y) of S a definite direction, namely, the direction of the line whose slope (dy/dx) is $f(x, y)$. We can think of these directions as indicated by short vectors drawn at the points.

To integrate equation (1) is to find a curve drawn in S, such that,

* Here, as in so many other cases in physics, a thorough knowledge of Linear Dependence is indispensable for an understanding of the subject in hand; *cf.* Bôcher, *Algebra*, Chaps. 3, 4, or better still, Bôcher, *Annals of Math.*, ser. 2, vol. 2 (1901), p. 81.

at each one of its points, it is tangent to the vector which pertains to that point. It would seem likely that such a curve exists through each point of S. For, if we start at any point (x_0, y_0) of S and go along the vector at that point to a point (x_1, y_1) near by; if from here we proceed along the vector pertaining to this latter point to a point (x_2, y_2) a little further on; and if we continue this process, we are thus led to a broken line, whose slope at any one of its points* differs but slightly from the value of $f(x, y)$ at that point. It is natural to expect this line to approach a certain curve as its limit when its vertices increase in number, the greatest distance between two successive vertices approaching 0. On introducing a suitable restriction on $f(x, y)$ (a reason for which will appear when we study singular solutions) it turns out that this is the case; *i.e.* that there is a curve,

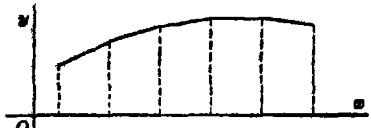


FIG. 83

$$y = \phi(x),$$

toward which all these broken lines converge, and that the slope of this curve at each point is that of the vector pertaining to this point. Analytically this means that the function $\phi(x)$ satisfies the given differential equation, or

$$\phi'(x) \equiv f[x, \phi(x)].$$

The condition to be imposed on $f(x, y)$ may be stated in the form that $f_y(x, y) = \partial f/\partial y$ shall exist and be continuous throughout S. This condition is somewhat more restrictive than is needed, but it includes the cases of importance which arise in practice. Moreover, when this condition is fulfilled, the solution is unique; *i.e.* the neighborhood of an arbitrary interior point of S is swept out just once by a one-parameter family of curves, no two of which have a point in common.

We note that the solution depends on an arbitrary constant, y_0. At first sight it might appear as if it depended on *two* arbitrary constants, x_0 and y_0. It does; and still there is only a one-parameter family of solutions involved, for we get all the solutions which course the neighborhood of the point (x_0, y_0) by holding x_0 fast and allowing y_0 alone to vary. For example, the right lines which have

* At a vertex, the slope of one of the lines abutting on it is just right, by construction. The slope of the other line is not far wrong.

a given slope ($\lambda = 2$, say) are given by the equation

$$y - y_0 = 2(x - x_0), \qquad \text{or} \qquad y = 2x + (y_0 - 2x_0).$$

Thus the two arbitrary constants x_0 and y_0 are together equivalent to but a single arbitrary constant,

$$b = y_0 - 2x_0.$$

Example. Consider the differential equation

$$\frac{dy}{dx} = r, \qquad r = \sqrt{x^2 + y^2}.$$

Let it be required to find approximately where the axis of x is cut by the solution which cuts the axis of y one unit above the origin.

The student should make an accurate drawing on squared paper, taking 10 cm. as the unit of length and making $x_{k+1} - x_k = \frac{1}{10}$ (*i.e.* 1 cm. long).

Simultaneous Differential Equations. A simultaneous system of the form

$$(2) \qquad \frac{dy}{dx} = F(x, y, z), \qquad \frac{dz}{dx} = \Phi(x, y, z)$$

can be treated in a similar manner. Let V be a region of space, at every point (x, y, z) of which the functions F and Φ are continuous. Draw through (x, y, z) a line whose direction components are 1, $F(x, y, z)$, $\Phi(x, y, z)$, and lay off a short vector along this line. A curve,

$$(3) \qquad y = f(x), \qquad z = \phi(x),$$

which, at each of its points, is tangent to the vector pertaining to that point, will represent a solution of the given system (2).

Starting at any point (x_0, y_0, z_0) of V, we can construct a broken line as in the earlier case, laying off first a short distance on the vector at (x_0, y_0, z_0). From the end, (x_1, y_1, z_1), of this line lay off a short distance on the vector pertaining to (x_1, y_1, z_1); and continue in this way. The broken line thus formed will approach a limiting curve, (3), which represents a solution of (2), provided $F(x, y, z)$ and $\Phi(x, y, z)$ admit first partial derivatives with respect to y and z, which are continuous throughout V.

The extension to the case of a simultaneous system of n equations in n dependent variables:

$$\frac{dy_1}{dx} = F_1(x, y_1, \cdots, y_n), \quad \cdots, \quad \frac{dy_n}{dx} = F_n(x, y_1, \cdots, y_n),$$

is immediate, and the existence theorem holds in that case, too. Such a system can be thrown into the equivalent form:

$$\frac{dx_1}{F_1(x_1, \cdots, x_n)} = \frac{dx_2}{F_2(x_1, \cdots, x_n)} = \cdots = \frac{dx_n}{F_n(x_1, \cdots, x_n)} = dt,$$

or
$$\frac{dx_1}{dt} = F_1(x_1, \cdots, x_n), \quad \cdot \quad \cdot \quad \cdot, \quad \frac{dx_n}{dt} = F_n(x_1, \cdots, x_n).$$

Through each point, $(x_1^0, \cdots, x_n^0)$ of the region $\mathfrak{R}$ of the n-dimensional space, in which the functions F_k are continuous, together with their partial derivatives of the first order, passes one and only one curve,

$$x_k = \phi_k(t, x_1^0, \cdots, x_n^0), \qquad k = 1, \cdots, n,$$

provided the functions F_k are not all zero at $(x_1^0, \cdots, x_n^0)$.

16. Continuation. Differential Equations of the Second Order, and Higher. Consider the differential equation

$$(1) \qquad \frac{d^2y}{dx^2} = F\left(x, y, \frac{dy}{dx}\right),$$

where $F(x, y, p)$ is continuous for all points (x, y, p) for which (x, y) lies in a given region S of the (x, y)-plane, and p is arbitrary. Through an arbitrary point (x, y) of S draw an arbitrary line, of slope p. Then (1) determines the circle of curvature of the solution of (1) which passes through (x, y) and has the slope p. Thus a small arc of the osculating circle at this point can be laid off, and we can proceed to build up an approximate solution by means of such arcs, much as in the earlier case, § 15.

A solution of (1) will be a function, $y = \phi(x)$, continuous together with its first and second derivatives, and such that its graph has its curvature at each point in agreement with (1), or

$$\pm \kappa = \frac{F(x, y, p)}{(1 + p^2)^{\frac{3}{2}}} \qquad p = \frac{dy}{dx}.$$

If, furthermore, the partial derivatives of F (of the first order) exist and are continuous, then through an arbitrary interior point (x_0, y_0) of S will pass a solution of (1) having an arbitrary slope, p_0; and the solution is thus uniquely determined near (x_0, y_0, p_0).

We see, moreover, that the solution depends on *two* arbitrary constants, which may be taken as y_0 and $p_0 = y_0'$.

Second Method. A second method of treating the differential equation (1) is the following. Introduce a new variable, $z = dy/dx$. Then (1) is replaced by the simultaneous system

$$(2) \qquad \frac{dy}{dx} = z, \qquad \frac{dz}{dx} = F(x, y, z).$$

If $F(x, y, z)$ is continuous, together with its partial derivatives of the first order, the system (2) will admit a unique solution passing through (x_0, y_0, z_0); § 15. Thus the existence theorem for (1) stated just above is here proved by means of the existence theorem of § 15.

The method can be extended to m simultaneous differential equations of orders $n_1, n_2, \cdots, n_m$ respectively in the dependent variables $y_1, \cdots, y_m$. They are seen to be equivalent to a system of $q = n_1 + n_2 + \cdots n_m$ simultaneous differential equations of the type

$$\frac{dz_i}{dx} = \Phi_i(x, z_1, \cdots, z_q), \qquad i = 1, 2, \cdots, q,$$

and their solution depends on q arbitrary constants.

17. Singular Solutions. Consider the differential equation

$$(1) \qquad \frac{dy^2}{dx^2} = 1 - y^2.$$

Nothing could be easier than to integrate it by our ordinary methods. First,

$$\frac{dy}{dx} = \pm\sqrt{1 - y^2}.$$

Next, separate the variables:

$$dx = \pm \frac{dy}{\sqrt{1 - y^2}}.$$

Hence
$$x = \pm \int \frac{dy}{\sqrt{1 - y^2}} = \mp \cos^{-1} y + c,$$

and so, finally,

$$(2) \qquad y = \cos(x - c),$$

— the "complete primitive," as the books call it, containing an arbitrary constant, and so comprising all the solutions of (1).

The only trouble with this result is that it is wrong. Not that the function (2) is not a solution for an arbitrary value of c, but in the assertion that *all* the solutions of (1) are given by (2) there is a blunder.

Let us study the differential equation from the point of view of § 15. Consider the two differential equations

$$(i)\quad \frac{dy}{dx} = \sqrt{1-y^2}; \qquad (ii)\quad \frac{dy}{dx} = -\sqrt{1-y^2}.$$

It is clear that any solution of (i) in a given interval $a \leqq x \leqq b$ is a solution of (1); and the same remark applies to (ii).

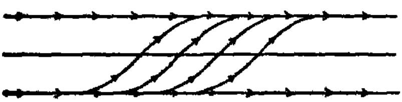


Fig. 84

Equation (i) assigns to each point (x, y) within the strip $-1 \leqq y \leqq 1$ of the (x, y)-plane a positive slope, and to every point on the boundary $(y = 1, y = -1)$ of the strip the slope 0. If (x_0, y_0) be an arbitrary interior point of the strip, there passes through it one and only one solution, and to the determination of this solution the analysis above applies:

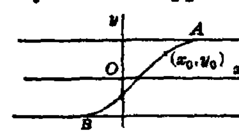


Fig. 85

$$\frac{dy}{dx} = \sqrt{1-y^2},$$

$$x = \int \frac{dy}{\sqrt{1-y^2}} = -\cos^{-1} y + C,$$

$$x_0 = -\cos^{-1} y_0 + C,$$

$$(3)\qquad x = -\cos^{-1} y + \cos^{-1} y_0 + x_0,$$

where each $\cos^{-1}$ means the principal value of the function (*Introduction to the Calculus*, Chap. VIII, p. 211).

The solution (3) proceeds forward till it meets the line $y = 1$ at A; and it runs backward till it touches the line $y = -1$ at B. *The continuation of this solution to the right of A is the function*

$$y = 1,$$

or the part of the *upper boundary of the strip* to the right of A. And likewise the solution is carried backward, to the left of B, by the function $y = -1$, or the part of the lower boundary of the strip to the left of B.

We see, then, that through every interior point of the strip passes one and only one integral curve of (i), and this solution is defined throughout the whole range of values for x, $-\infty < x < +\infty$.

Not so, however, with a point (x_0, y_0) on the boundary of the strip. Suppose $y_0 = -1$. Then that part of any solution passing through (x_0, y_0) which lies to the left of x_0 is uniquely determined; it is $y = -1$. But to the right we may proceed along this same line forever, thus having the solution $y=-1$, $-\infty < x < +\infty$; or we may leave the line at any point $x_1 \geqq x_0$, pass along the curve (3) to the other boundary, $y=1$, and then continue forevermore on this line.

Thus we see that through any point on the boundary of the strip pass infinitely many solutions of (*i*).

Precisely similar results hold for (*ii*). The solution passing through a point (x_0, y_0) within the strip is given by

(4) $$x = \cos^{-1} y - \cos^{-1} y_0 + x_0,$$

where, as before, the principal value of each $\cos^{-1}$ is meant. And this solution continues along the boundary.

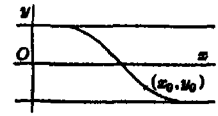


FIG. 86

The Solution of (1). We see now how to put together other solutions of (1) than those given by (2). First, the solutions of (*i*) and (*ii*) just discussed are all solutions of (1). Secondly, we may start with the arc AB of a solution of (*i*), proceed to the right of A an arbitrary distance, switch to a solution of (*ii*), follow the line $y=-1$, as far as we like, then switch to a solution of (*i*); and so on.

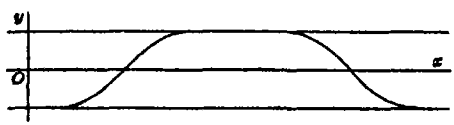


FIG. 87

Do we get, even in this way, all the solutions of (1)? For an interior point (x_0, y_0) any solution of (1) is given either by (3) or by (4) till it reaches the boundary. For a boundary point (x_0, y_0) any solution, considered to the right of the point, either coincides with the boundary for an interval, or it has points distinct from the boundary in every neighborhood to the right of x_0. In the latter case, it must switch to a branch (3) or (4) to the right of x_0, and the transition must obviously be made at the point x_0. Similarly for the left-hand neighborhood of a boundary point. The solution of (1) is now complete.

Example from Physics. Consider a simple pendulum, *Introduction to the Calculus*, p. 373. The differential equation of the first order* is

(5) $$\frac{d\theta^2}{dt^2} = n^2(\alpha^2 - \theta^2), \qquad n = \sqrt{\frac{g}{l}},$$

* This is the approximate equation for small arcs; but the reasoning applies equally well to the accurate equation, l. c. (2).

and can be written down at once from the principle of Work and Energy. Suppose that the pendulum gets stuck when it reaches its highest point, and lodges there till some one releases it. This may happen each time it comes to its highest point, and each time it may remain at rest for an arbitrary interval of time. The equation of Work and Energy is the same for this case as for the case ordinarily considered, namely, (5).

From this it appears that (5) regarded as the mathematical formulation of the problem of Simple Pendulum Motion, is not adequate, since (5) admits other solutions, too. The same remark applies to many of the deductions given in Mechanics, which are based on the principle of energy and operate with differential equations of the first order, which are not linear. On the other hand, this situation cannot arise when the solution is based on Newton's Second Law of Motion and the formulation

$$\frac{d^2\theta}{dt^2} = -\frac{g}{l}\theta. \qquad \left(\text{or} = -\frac{g}{l}\sin\theta.\right)$$

This differential equation has only one solution, and that, the solution of the problem.

18. Continuation. The General Case. The central fact illustrated by the example of § 17 may be stated as follows. The family of solutions (2) have an *envelope*, namely, the lines $y = 1$ and $y = -1$. An arc of the envelope (a segment of either line) obviously must also yield a solution of the differential equation; and yet this solution is not contained in those given by (2). Such a solution is called a *singular solution*.

We can generalize and say: Let a differential equation of the first order be satisfied by a family of curves,

$$(1) \qquad y = \phi(x, c),$$

and let these curves all be tangent to a curve

$$(2) \qquad y = \psi(x).$$

Through any point (x_0, y_0) passes a curve, $y = \phi(x, c_0)$, of the family (1); and the function $\phi(x, c_0)$ satisfies the differential equation in the neighborhood of the point $x = x_0$. If the curve (2) is not contained in the family (1), it is called a *singular solution*.

For example, consider the differential equation

$$\frac{dy}{dx} = 3y^{\frac{2}{3}}.$$

A family of solutions is seen to be

$$y = (x - c)^3.$$

These curves are all tangent to the axis of x, and this line, $y = 0$, is seen to be a solution of the given differential equation. But it is not one of the above family; it is a singular solution.

Clairaut's Equation. Consider the differential equation

$$y = x\frac{dy}{dx} + f\left(\frac{dy}{dx}\right), \tag{3}$$

where $f(p)$ is continuous, together with $f'(p)$ and $f''(p)$, and $f''(p) \neq 0$.

The "general solution" (1) of (3) can be written down at sight:

$$y = cx + f(c), \tag{4}$$

where c is an arbitrary constant. Thus we have a family of straight lines.

This family, however, has an envelope; for, differentiate (4) partially with respect to c, Chap. VIII, § 1:

$$0 = x + f'(c).$$

Thus the envelope is defined by the equations

$$\begin{cases} x = -f'(c), \\ y = -cf'(c) + f(c). \end{cases} \tag{5}$$

This curve represents a *singular solution.*

Example. $$y = x\frac{dy}{dx} - \frac{1}{2}\left(\frac{dy}{dx}\right)^2.$$

Here, the general solution corresponds to the straight lines

$$y = cx - \tfrac{1}{2}c^2.$$

The singular solution is:

$$x = c, \qquad y = \tfrac{1}{2}c^2,$$

or the parabola:

$$2y = x^2.$$

IV. Solution by Series. Integrating Factor

19. Bessel's Functions. Zonal Harmonics. The problems of Mathematical Physics lead to certain homogeneous linear differential equations of the second order with variable coefficients which are very simple functions. The most important equations of this class are:

(*a*) **Bessel's Equation:**

$$\frac{d^2y}{dx^2}+\frac{1}{x}\frac{dy}{dx}+\left(1-\frac{n^2}{x^2}\right)y=0;$$

(*b*) Legendre's Equation:

$$\frac{d}{dx}\left\{(1-x^2)\frac{dy}{dx}\right\}+m(m+1)y=0.$$

The first of these cannot be solved in terms of the elementary functions. It defines a new class of functions, the Bessel's Functions of the First Kind, denoted by $J_n(x)$, and those of the Second Kind, denoted by $K_n(x)$.

The Series for $J_0(x)$. On setting $n=0$, Bessel's Equation becomes:

(1) $$\frac{d^2y}{dx^2}+\frac{1}{x}\frac{dy}{dx}+y=0.$$

Let us see if we can obtain a solution of this differential equation in the form of a power series,

(2) $$y=a_0+a_1x+a_2x^2+\cdots.$$

Writing (1) in the form

(3) $$x\frac{d^2y}{dx^2}+\frac{dy}{dx}+xy=0,$$

we compute the left-hand side of (3) by means of (2):*

$$x\frac{d^2y}{dx^2}= \quad 2\cdot 1\,a_2x+3\cdot 2\,a_3x^2+\cdots+(n+2)(n+1)a_{n+2}x^{n+1}+\cdots$$

$$\frac{dy}{dx}=a_1+2\,a_2x+3\,a_3x^2+\cdots+(n+2)\,a_{n+2}x^{n+1}+\cdots$$

$$xy=a_0x+a_1x^2+\cdots+a_nx^{n+1}+\cdots$$

On adding these three equations together we obtain a single power series in x, whose constant term is a_1. The coefficients of all subsequent terms are given by the formula:

$$(n+2)(n+1)\,a_{n+2}+(n+2)\,a_{n+2}+a_n=(n+2)^2a_{n+2}+a_n.$$

Set each of these coefficients equal to 0; thus

(4) $$a_1=0, \qquad a_{n+2}=-\frac{1}{(n+2)^2}a_n.$$

* We assume here without proof that a power series can be differentiated term-by-term, as if it were a polynomial. — The object in writing the general term as the one in x^{n+1}, rather than as the one in x^n, is to obtain a somewhat simpler form of the relation between the coefficients of (2).

The first coefficient in (2), namely, a_0, is arbitrary. From (4) we find:

$$a_2 = -\frac{1}{2^2}a_0, \qquad a_4 = \frac{1}{2^2\,4^2}a_0, \qquad a_6 = -\frac{1}{2^2\,4^2\,6^2}a_0, \cdots.$$

Furthermore, since $a_1 = 0$, it follows from (4) that $a_3 = 0$, $a_5 = 0$, etc.

Each a_n is thus seen to contain a_0 as a factor, and since we do not care particularly what constant factor is multiplied into a series (2) which yields a solution, we will set $a_0 = 1$. (2) then becomes the series which defines the function known as $J_0(x)$:

(5) $$J_0(x) = 1 - \frac{x^2}{2^2} + \frac{x^4}{2^2\,4^2} - \frac{x^6}{2^2\,4^2\,6^2} + \cdots.$$

This series converges for all values of x, and it converges rapidly. The series for $J_n(x)$ and $K_n(x)$ will be found in Peirce's *Table of Integrals*, p. 87. They can be verified by direct substitution in the above differential equation (a). When n is not an integer, $J_n(x)$ and $J_{-n}(x)$ afford two linearly independent solutions of (a), and there is no need of introducing a function $K_n(x)$ (which in this case is not defined). But if n is an integer, $J_n(x)$ and $J_{-n}(x)$ become linearly dependent, and $K_n(x)$ is needed to furnish a second solution.

Zonal Harmonics. Legendre's Equation, (b), can be treated in a precisely similar manner. If a solution is assumed in the form of a power series,

$$y = a_0 + a_1x + a_2x^2 + \cdots,$$

it is found that the relation

$$(n+1)[(n+2)a_{n+2} - na_n] + m(m+1)a_n = 0,$$

or

$$a_{n+2} = \frac{n(n+1) - m(m+1)}{(n+1)(n+2)}a_n$$

holds for $n = 0, 1, 2, \cdots$. The coefficients a_0 and a_1 are arbitrary, and we get two linearly independent solutions by setting first one of these coefficients, and then the other, equal to zero.

When m is a positive integer, or zero, one of these solutions reduces to a polynomial. For the coefficients a_{m+2}, a_{m+4}, $\cdots$ are seen to vanish, and thus one of the solutions breaks off with the term a_mx^m. The other solution is not a polynomial.

Let the polynomial solution be arranged according to descending powers of x:

$$a_mx^m + a_{m-2}x^{m-2} + \cdots.$$

The coefficients can be computed in terms of a_m by reversing the last formula; thus

$$a_{m-2} = -\frac{m(m-1)}{2(2m-1)}a_m,$$

$$a_{m-4} = \frac{m(m-1)(m-2)(m-3)}{2\cdot 4\cdot(2m-1)(2m-3)}a_m, \cdots.$$

It turns out to be convenient to choose as a_m the number

$$a_m = \frac{(2m-1)(2m-3)\cdots\cdot 3\cdot 1}{m!}.$$

The polynomial thus arising is known as a *Zonal Harmonic* or a *Legendre's Coefficient*, and is written as $P_m(x)$; *cf.* Chap. XVI, §§ 5, 6:

$$P_m(x) = \frac{(2m-1)(2m-3)\cdots\cdot 3\cdot 1}{m!}\left\{x^m - \frac{m(m-1)}{2(2m-1)}x^{m-2}\right.$$

$$\left.+\frac{m(m-1)(m-2)(m-3)}{2\cdot 4\cdot(2m-1)(2m-3)}x^{m-4} - \cdots\right\}.$$

In particular,

$$P_0 = 1, \qquad P_1 = x, \qquad P_2 = \tfrac{1}{2}(3x^2-1), \qquad P_3 = \tfrac{1}{2}(5x^3-3x).$$

20. Integrating Factor. Let $M = f(x, y)$ and $N = \phi(x, y)$ be two functions which, together with their partial derivatives of the first order (in particular, $\partial M/\partial y$ and $\partial N/\partial x$) are continuous throughout a region S of the (x, y)-plane. The expression

(1) $$M\,dx + N\,dy$$

will not in general be the differential of any function $u = F(x, y)$, since for this to be the case we must have $\partial M/\partial y = \partial N/\partial x$; cf. Chap. XI, § 7.

It is, however, conceivable that, on multiplying (1) by a suitable factor, $\rho = \omega(x, y)$, the product

(2) $$\rho(M\,dx + N\,dy)$$

may become an "exact differential":*

$$du = \rho M\,dx + \rho N\,dy.$$

* The following treatment presupposes entire familiarity with the developments of Chap. XI, §§ 1–7. This is not a technicality. There is no short cut to the integrating factor, whereby an understanding of the subject matter of those paragraphs can be avoided.

That such a function, $u = F(x, y)$, exist, it is both necessary and sufficient that

$$\frac{\partial F}{\partial x}dx + \frac{\partial F}{\partial y}dy = \rho M dx + \rho N dy,$$

or

$$(3) \qquad F_x = \rho M, \qquad F_y = \rho N.$$

A pair of functions, F and ρ, satisfying this last pair of equations can be found as follows.

Let (x_0, y_0) be a point of S, in which not both of the functions M and N vanish. Suppose $N(x_0, y_0) \neq 0$. Then the differential equation

$$(4) \qquad \frac{dy}{dx} = -\frac{M}{N}$$

admits a one-parameter family of solutions which course the neighborhood of this point (cf. § 15), and which can be expressed in the form:

$$(5) \qquad F(x, y) = C,$$

where F is continuous, together with its first partial derivatives, and $F_y \neq 0$.* The slope of the curve (5) at an arbitrary point (x, y) is

$$\frac{dy}{dx} = -\frac{F_x}{F_y}.$$

But dy/dx is also given by (4). Hence at every point (x, y) of S we have

$$(6) \qquad -\frac{F_x}{F_y} = -\frac{M}{N}.$$

Let

$$(7) \qquad \rho = \frac{F_y}{N}, \qquad \text{or} \qquad F_y = \rho N.$$

It follows, then, from (6) that $F_x = \rho M$, and the proof is complete.

The factor ρ is known as an *integrating factor*. There is an infinite number of such factors. Thus in the case of

$$y\,dx - x\,dy$$

it is evident by inspection that

$$\frac{1}{x^2}, \qquad \frac{1}{y^2}, \qquad \frac{1}{x^2 + y^2}$$

are all integrating factors.

* The proof of this existence theorem is given in the treatment of the theory of differential equations.

Let u be the function corresponding to the particular integrating factor ρ; let σ be any second integrating factor, and let v be the function corresponding to it:

$$dv = \sigma M\,dx + \sigma N\,dy.$$

Then
$$\frac{\partial u}{\partial x} = \rho M, \qquad \frac{\partial u}{\partial y} = \rho N;$$
$$\frac{\partial v}{\partial x} = \sigma M, \qquad \frac{\partial v}{\partial y} = \sigma N.$$

Hence
$$\frac{\partial (u, v)}{\partial (x, y)} \equiv 0$$

and
$$v = \Omega(u).$$

We now infer that
$$\sigma = \Omega'(u)\,\rho.$$

For
$$\frac{\partial v}{\partial y} = \Omega'(u)\frac{\partial u}{\partial y}, \qquad \text{or} \qquad \sigma N = \Omega'(u)\,\rho N,$$

and $N \neq 0$.

Conversely, let $f(u)$ be any continuous function of u. Then

$$\sigma = f(u)\,\rho$$

will be an integrating factor. For

$$\int\limits_{(a,b)}^{(x,y)} \sigma M\,dx + \sigma N\,dy = \int\limits_{(a,b)}^{(x,y)} f(u)\,\{\rho M\,dx + \rho N\,dy\} = \int\limits_{C} f(u)\frac{du}{ds}\,ds,$$

and thus the first integral is independent of the path in any simply connected region. It defines, therefore, a function v, and

$$\frac{\partial v}{\partial x} = \sigma M, \qquad \frac{\partial v}{\partial y} = \sigma N, \qquad\qquad v = F(u) = \int f(u)\,du.$$

Three Variables: $P\,dx + Q\,dy + R\,dz$. That an integrating factor, ρ, should exist in this case it is necessary and sufficient that

$$\frac{\partial(\rho R)}{\partial y} = \frac{\partial(\rho Q)}{\partial z}, \qquad \frac{\partial(\rho P)}{\partial z} = \frac{\partial(\rho R)}{\partial x}, \qquad \frac{\partial(\rho Q)}{\partial x} = \frac{\partial(\rho P)}{\partial y}.$$

These conditions cannot in general be fulfilled. It is readily seen that a necessary condition for the existence of an integrating factor is:

(8) $$P\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right) + Q\left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right) + R\left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) = 0.$$

This condition can be shown, conversely, to be sufficient; Goursat, *Cours d'analyse mathématique*, vol. 2, § 442.

EXERCISES

1. Show that all the integrating factors of

$$y\,dx - x\,dy$$

are given by $f\left(\frac{y}{x}\right)\cdot\frac{1}{x^2}$, where $f(u)$ is any continuous function of u.

2. Determine all the integrating factors of

$$(2y + \cos x)\,dx - dy.$$

3. The same question for

$$(7x - 5y)\,dx + (3x + 2y)\,dy.$$

V. Partial Differential Equations

21. Nature of the Solution. The simplest partial differential equation one can well imagine is

$$\frac{\partial u}{\partial x} = 0, \tag{1}$$

where u is a function of the two independent variables (x, y). Its most general solution can be written down at sight:

$$u = f(y), \tag{2}$$

where $f(y)$ is *any* function of y whatsoever, continuous or discontinuous — even discontinuous for every value of y.*

A further example is the partial differential equation:

$$\frac{\partial^2 u}{\partial x\,\partial y} = 0. \tag{3}$$

Since this differential equation says that

$$\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial y}\right) = 0,$$

it is seen to come under the case just considered under (1), and so

$$\frac{\partial u}{\partial y} = f(y) \tag{4}$$

is a first integral, $f(y)$ being arbitrary.

But here we meet with a difficulty, for we cannot go further with a function $f(y)$ which cannot be integrated. We are thus compelled

* For example, $f(y)$ might be $= 0$ when y is a rational number, and $= 1$ when y is irrational.

here to restrict the function $f(y)$ at least to being integrable; we will require that it be continuous. Let, then,

$$F(y) = \int f(y)\,dy.$$

And now

$$u = F(y) + \Phi(x), \tag{5}$$

where $\Phi(x)$ is wholly arbitrary.

Here, however, we meet still another difficulty, for if $\Phi(x)$ cannot be differentiated, then

$$\frac{\partial^2 u}{\partial y\,\partial x} \neq \frac{\partial^2 u}{\partial x\,\partial y},$$

and we are thus led to distinctions which it is embarrassing to have to make.

In order that our attention may not be distracted at the outset by details which obscure the things of first importance, we will agree to consider only such solutions of partial differential equations of the first (second) order as are continuous, together with their first (first and second) partial derivatives, throughout a region S of the (x, y)-plane; or at least throughout each of a set of regions, S_1, S_2, $\cdots$, into which the given region S can be cut up.

Thus we should demand that the functions $F(y)$ and $\Phi(x)$ in (5) be continuous together with their first and second * derivatives, except for isolated values of y and x, or along certain curves.

It will be observed, in the foregoing examples, that the solution of a partial differential equation of the first order involves *one arbitrary function*, and the solution of one of the second order involves *two arbitrary functions*. This is typical for the general case, and is analogous to the fact that the solution of an ordinary differential equation of the first order involves one arbitrary constant, the solution of one of the second order, two arbitrary constants.

22. Linear Partial Differential Equations of the First Order. Consider the differential equation

$$A\frac{\partial z}{\partial x} + B\frac{\partial z}{\partial y} = C, \tag{1}$$

where A, B, C are any three continuous functions of (x, y, z)

* It happens that, for this particular differential equation, no assumption about the second derivatives is needed. But in the transformations of (3) considered below, the second derivatives enter.

throughout a region V of space, A and B not vanishing simultaneously. Let

(2) $$z = F(x, y)$$

be a solution of (1). The direction components of the normal to the surface (2) at any point are: $\partial z/\partial x$, $\partial z/\partial y$, -1. On the other hand, to each point of V corresponds a definite direction whose direction components are A, B, C. And now the given differential equation says that the latter direction shall always lie in the tangent plane to (2).

This geometric fact gives us a hint as to how to integrate (1). The geometric picture of a region of space, to each point of which a direction is assigned, is one we have met before. It suggests providing each point with a little vector drawn in the prescribed direction, and then seeking a two-parameter family of curves that just fill out the region, each curve being tangent at every point to the vector pertaining to that point. These curves will be defined analytically by the simultaneous system of ordinary differential equations,

(3) $$\frac{dy}{dx} = \frac{B}{A}, \qquad \frac{dz}{dx} = \frac{C}{A},$$

or $$\frac{dx}{A} = \frac{dy}{B} = \frac{dz}{C}.$$

Their equations, therefore, can be written in either of the forms

(4) $$y = f(x, x_0, y_0, z_0), \qquad z = \phi(x, x_0, y_0, z_0);$$

(4′) $$x = f_1(t, x_0, y_0, z_0), \qquad y = f_2(t, x_0, y_0, z_0), \qquad z = f_3(t, x_0, y_0, z_0).$$

These curves are known as the *characteristics* of the differential equation (1).

Let Γ be a curve (open or closed) drawn in V and not coinciding along any arc with any of the curves (4) or (4′). Through each point (x_0, y_0, z_0) of Γ passes a curve (4) or (4′), and the one-parameter family of curves thus obtained forms a surface,

(5) $$z = \Phi(x, y),$$

as will be shown presently.

This surface (5) is an integral surface of (1). For, its normal at any point is perpendicular to the particular curve (4) or (4′) through that point. The direction components of the curve are A, B, C; those of the normal are Φ_x, Φ_y, -1. Hence (1) is satisfied.

Analytically this surface is expressed as follows. Let Γ be given in parametric form by the equations

(6) $$x = \psi(u), \qquad y = \chi(u), \qquad z = \omega(u).$$

And now substitute in (4′) these three functions for x_0, y_0, z_0 respectively:

$$x = f_1(t, \psi(u), \chi(u), \omega(u)), \qquad y = f_2(t, \psi(u), \chi(u), \omega(u)), \qquad z = \text{etc.}$$

Conversely, any solution of (1) yields a surface $z = F(x, y)$ which is swept out by a one-parameter family of curves (4′). For, through each of its points passes a curve (4′), and any such curve lies wholly in the surface, as the student can readily prove for himself.

We observe that the general solution of (1) depends on *one arbitrary function.* At first sight there seem to be two (or even three) such functions, corresponding to Γ. But if we cut an arbitrary solution of (1) by a plane, this plane curve is sufficient to define all the solutions near the given one, and a plane curve is equivalent to a single arbitrary function.

Proof of (5). In (4) hold x_0 fast, and for simplicity let $x_0 = 0$. Rewrite (4):

(7) $$y = g(x, y_0, z_0), \qquad z = h(x, y_0, z_0).$$

Then

(8) $$g(0, y_0, z_0) = y_0, \qquad h(0, y_0, z_0) = z_0.$$

Hence, when $x = 0$,

(9) $$\frac{\partial y}{\partial y_0} = 1, \qquad \frac{\partial y}{\partial z_0} = 0, \qquad \frac{\partial z}{\partial y_0} = 0, \qquad \frac{\partial z}{\partial z_0} = 1.$$

Let the curve Γ be given as follows:

(10) $$z_0 = \omega(y_0), \qquad x = x_0 = 0,$$

where ω is continuous, together with its first derivative. On substituting for z_0 in (7) the value $\omega(y_0)$, it is seen that the first equation can be solved for y_0:

(11) $$y_0 = \chi(x, y),$$

where χ is continuous, together with its first derivatives. Eliminate y_0 and z_0 from the second equation (7) by means of (10) and (11); thus (5) results.

In equations (4′), regarded as applying to the neighborhood of a point (x_0, y_0, z_0), the parameter t can always be taken as x, when $A \neq 0$, and as y, when $B \neq 0$, at that point. Thus the foregoing proof is general if $A \neq 0$; otherwise all that is necessary is to allow x and y to interchange their rôles.

23. General Partial Differential Equation of the First Order. Consider the partial differential equation

$$(1) \qquad F(x, y, z, p, q) = 0, \qquad p = \frac{\partial z}{\partial x}, \qquad q = \frac{\partial z}{\partial y},$$

where F is continuous, together with its first partial derivatives. We assume that (x, y, z) is a point of a region V of space, that p and q may have any values whatever, and that F_p and F_q do not vanish simultaneously.

Let (x_0, y_0, z_0) be held fast, and consider the lines through this point, whose direction components, p, q, -1, satisfy (1). These lines sweep out a cone (N), and their normal planes through (x_0, y_0, z_0):

$$(2) \qquad z - z_0 = p(x - x_0) + q(y - y_0),$$

envelop a cone (T), whose generators are determined by (2) and the further equation (Chap. VIII, § 4, Ex. 2):

$$0 = \frac{\partial p}{\partial q}(x - x_0) + (y - y_0)$$

or

$$(3) \qquad \frac{x - x_0}{P} = \frac{y - y_0}{Q},$$

the notation here and later being:

$$\frac{\partial F}{\partial x} = X, \qquad \frac{\partial F}{\partial y} = Y, \qquad \frac{\partial F}{\partial z} = Z, \qquad \frac{\partial F}{\partial p} = P, \qquad \frac{\partial F}{\partial q} = Q.$$

Consider now a solution of (1),

$$(4) \qquad z = \Phi(x, y), \qquad \text{the surface } S.$$

The tangent plane to S at an arbitrary point (x_0, y_0, z_0) is given by (2), and the generator of (T) which lies in that plane lies also in the plane (3). Thus a direction is determined at every point of S, and these directions can be visualized by short vectors, which may be curved so as to lie actually in S.*

* We may think of the cone (T), roughly speaking, as *tangent* to S at (x_0, y_0, z_0) along the direction above determined. More precisely: let an arbitrary curve, Γ, be drawn on S through (x_0, y_0, z_0), and let P be a neighboring point of Γ. Consider the distance, ζ, from P to the cone (T). If Γ have, in particular, the above direction, then ζ will be an infinitesimal of higher order than if Γ has a different direction.

We have here a geometric picture closely similar to that of § 15, and we should expect to find the surface S swept out by a one-parameter family of curves, each of which is tangent at every one of its points to the direction pertaining to that point. This is precisely what happens.

The curves C of this family are determined as follows.* The differential equation

$$(5) \qquad \frac{dx}{P} = \frac{dy}{Q},$$

where z is given by (4), has as its solution in the neighborhood of the point (x_0, y_0) a one-parameter family of curves which sweep out this neighborhood just once. And now the cylinders on these curves as directrices, with their elements parallel to the axis of z, cut out from S the curves C, and these sweep out the part of S that lies in the neighborhood of (x_0, y_0, z_0) just once.

For convenience, introduce a parameter, t, setting each side of (5) equal to dt. Thus

$$dx = P dt, \qquad dy = Q dt.$$

Furthermore, along any one of these curves C, we have

$$dz = p\,dx + q\,dy = (pP + qQ)\,dt,$$

and so (5) can be extended to read:

$$(6) \qquad \frac{dx}{P} = \frac{dy}{Q} = \frac{dz}{pP + qQ} = dt.$$

We can go further and compute dp and dq along the curve C. We have:†

$$dp = r\,dx + s\,dy, \qquad dq = s\,dx + t\,dy,$$

or

$$(7) \qquad dp = (rP + sQ)\,dt, \qquad dq = (sP + tQ)\,dt,$$

where $r = \frac{\partial^2 z}{\partial x^2} = \frac{\partial p}{\partial x}, \qquad s = \frac{\partial^2 z}{\partial x \partial y} = \frac{\partial p}{\partial y} = \frac{\partial q}{\partial x}, \qquad t = \frac{\partial^2 z}{\partial y^2} = \frac{\partial q}{\partial y}.$

* It is assumed that (x_0, y_0, z_0) is an interior point of V; that $\Phi(x, y)$ is continuous, together with its partial derivatives of the first order, throughout the neighborhood of the point (x_0, y_0); that $z_0 = \Phi(x_0, y_0)$, $p_0 = \Phi_x(x_0, y_0)$, $q_0 = \Phi_y(x_0, y_0)$; and that the partial derivatives of F, not only of the first, but also of the second order, are continuous functions of the five arguments in the neighborhood of $(x_0, y_0, z_0, p_0, q_0)$. Finally, let (Chap. VIII, § 4, Ex. 2)

$$F_{qq} P^2 - 2 F_{pq} PQ + F_{pp} Q^2 \neq 0.$$

† It happens that the letter t is used here in two senses; but no confusion will arise.

On the other hand, equation (1) is satisfied identically in x and y by the function (4). Hence, differentiating (1) partially with respect to x and y successively, we have:

$$
(8) \qquad \begin{cases} X + pZ + rP + sQ = 0, \\ Y + qZ + sP + tQ = 0. \end{cases}
$$

From (7) and (8) it appears that

$$
X + pZ + \frac{dp}{dt} = 0, \qquad\qquad Y + qZ + \frac{dq}{dt} = 0.
$$

Thus equations (6) can be extended once again and we have:

$$
(9) \qquad \frac{dx}{P} = \frac{dy}{Q} = \frac{dz}{pP + qQ} = \frac{-dp}{X + pZ} = \frac{-dq}{Y + qZ} = dt.
$$

This is a system of five total differential equations in the five dependent variables x, y, z, p, q and the independent variable t. They define a family of curves in the five-dimensional space of the variables x, y, z, p, q:

$$
(10) \qquad \begin{cases} x = f_1(t;\ x_0, y_0, z_0, p_0, q_0), \\ y = f_2(\quad `` \quad ), \\ z = f_3( \quad `` \quad), \\ p = f_4(\quad `` \quad ), \\ q = f_5( \quad `` \quad). \end{cases}
$$

It is assumed that $t = 0$ corresponds to $x = x_0$, $y = y_0, \cdots, q = q_0$.

Characteristics. Equations (10) define what is known as the family of *characteristics* of the given differential equation. The first three of these equations determine a curve C through (x_0, y_0, z_0), and the last two assign to each of its points a definite normal (with the direction-components $p, q, -1$) or a definite tangent plane perpendicular to this normal. We can combine both these conceptions into one composite picture by thinking of the curve C as embedded in a narrow strip of surface such that, at each point of C, the tangent plane to the surface has the orientation belonging to that point. Such a strip is called a *characteristic strip.* We shall say that it *lies in* the surface S, and we can think of S as made up of a one-parameter family of such characteristic strips.

Thus far, a definite solution (4) of (1) has underlain the entire treatment. We will begin now at the other end, starting with an entirely arbitrary family of characteristic strips, and inquire how the most general solution of (1), of the type (4), can be built up out of them.

24. Integration by Characteristics. Let $\mathfrak{A}:(a, b, c, \alpha, \beta)$ be any point in the five-dimensional space of the variables (x, y, z, p, q) such that $A:(a, b, c)$ is an interior point of V and $F(x, y, z, p, q)$ is continuous, together with its partial derivatives of the first and second orders, throughout the neighborhood of $\mathfrak{A}$. Moreover,

$$F(a, b, c, \alpha, \beta)=0, \quad \text{but} \quad P=F_p, \quad Q=F_q \quad \text{not both 0 in } \mathfrak{A}.$$

Then the system of differential equations (9) defines a four-parameter family of curves given by (10), which sweep out the neighborhood of $\mathfrak{A}$ just once, $(x_0, y_0, z_0, p_0, q_0)$ being an arbitrary point of this neighborhood.

Let C be the curve in V which is represented by the first three equations (10) and, in particular, let $C_{\mathfrak{A}}$ be the curve C which corresponds to the initial values $\mathfrak{A}$. Then any solution of (1),

$$(11) \qquad z=\Psi(x, y),$$

where $\Psi(x, y)$ is continuous, together with its partial derivatives of the first order, throughout the *entire* neighborhood of the point (a, b) and

$$c=\Psi(a, b), \qquad \alpha=\Psi_x(a, b), \qquad \beta=\Psi_y(a, b),$$

must contain the curve $C_{\mathfrak{A}}$ and the corresponding characteristic strip. For these are uniquely determined by the initial values corresponding to $\mathfrak{A}$.

Let $T_{\mathfrak{A}}$ be a plane through A, whose normal has the direction-components $(\alpha, \beta, -1)$. Then $C_{\mathfrak{A}}$ is tangent to $T_{\mathfrak{A}}$ at A.

Let D be any curve through A tangent to $T_{\mathfrak{A}}$ there, but not tangent to $C_{\mathfrak{A}}$:

$$D: \qquad x=\psi_1(u), \qquad y=\psi_2(u), \qquad z=\psi_3(u),$$

where $\psi_k(u)$ is continuous, together with its first derivative, and not all the $\psi_k'(0)$'s vanish; moreover,

$$(12) \qquad \psi_3'(0)=\alpha\,\psi_1'(0)+\beta\,\psi_2'(0).$$

Hence $\psi_1'(0)$ and $\psi_2'(0)$ are not both 0.

We can now state the fundamental existence theorem relating to (1).

Existence Theorem. *There exists one and only one solution of* (1):

$$F(x, y, z, p, q) = 0,$$

which satisfies the conditions imposed on the function (11) *and is such that the surface represented by* (11) *contains the curve D.*

In order to prove this theorem we show how a one-parameter family of characteristic strips may be picked out of (10) by means of D so as to sweep out the solution in question.

Determination of p_0, q_0. If a solution such as is demanded by the theorem is to exist, then the values of p and q along D, or p_0 and q_0, must satisfy the two conditions:

$$(13) \qquad \begin{cases} F[\psi_1(u),\ \psi_2(u),\ \psi_3(u),\ p_0, q_0] = 0, \\ \psi_1'(u)\, p_0 + \psi_2'(u)\, q_0 = \psi_3'(u), \end{cases}$$

the second being obtained by observing that (11) is satisfied identically along D, and that, at any one of these points,

$$\Psi_x[\psi_1(u),\ \psi_2(u)] = p_0, \qquad \Psi_y[\psi_1(u),\ \psi_2(u)] = q_0.$$

Can equations (13) be solved, however, for p_0, q_0? This question is answered in the affirmative by the Implicit Function Theorem, Chap. V, § 12. For, first, the above equations hold by hypothesis when $u = 0$ and $p_0 = \alpha$, $q_0 = \beta$. Secondly, the Jacobian of the functions F and

$$H = \psi_3' - p_0\psi_1' - q_0\psi_2',$$

namely,

$$\frac{\partial(F, H)}{\partial(p_0, q_0)} = -\begin{vmatrix} P_0 & Q_0 \\ \psi_1' & \psi_2' \end{vmatrix} = -\left(\frac{\partial x}{\partial t}\psi_2' - \frac{\partial y}{\partial t}\psi_1'\right)_{t=0,\ u=0}$$

is not zero, since its vanishing would mean that the projections of $C_{\mathfrak{A}}$ and D on the (x, y)-plane are tangent at (a, b).

Denote these functions as follows:

$$p_0 = \psi_4(u), \qquad q_0 = \psi_5(u).$$

These five functions, $\psi_k(u)$, then, are the values which $x_0, \cdots, q_0$, shall have in (10). The first three of them are as general as the curve D; the last two are a direct consequence of the choice of the first three.

Restatement of the Conditions for $\psi_k(u)$. It is useful to restate the conditions imposed on all five functions $\psi_k(u)$ independently of the curve D. Let

$$(14) \quad x_0 = \psi_1(u), \quad y_0 = \psi_2(u), \quad z_0 = \psi_3(u), \quad p_0 = \psi_4(u), \quad q_0 = \psi_5(u)$$

be continuous, together with the first derivative, in the neighborhood of the point $u=0$ and take on respectively the values a, b, c, α, β there. Let

$$(15)\quad \begin{cases} (i) & F(x_0, y_0, z_0, p_0, q_0)=0, & u=u; \\ (ii) & \dfrac{\partial z}{\partial u}=p\dfrac{\partial x}{\partial u}+q\dfrac{\partial y}{\partial u}, & t=0, \quad u=u; \\ (iii) & \begin{vmatrix} P & Q \\ \dfrac{\partial x}{\partial u} & \dfrac{\partial y}{\partial u} \end{vmatrix} \neq 0, & t=0, \quad u=0 \end{cases}$$

In (*ii*) and (*iii*) $x, y, \cdots, q$ are given by (10), the functions $x_0, \cdots, q_0$ being replaced in f_k by $\psi_1(u), \cdots, \psi_5(u)$.

The theorem then is that the first three equations (10) define a surface which represents a solution (11) of (1).

Proof of the Theorem. We observe that equations (10), in which the substitution (14) has been made, give:

$$(16)\qquad \frac{\partial x}{\partial t}=P, \quad \cdots, \quad \frac{\partial q}{\partial t}=-Y-qZ, \qquad \text{by (9)};$$

$$(17)\qquad \left.\frac{\partial x}{\partial u}\right|_{t=0}=\psi_1'(u), \quad \cdots, \quad \left.\frac{\partial q}{\partial u}\right|_{t=0}=\psi_5'(u).$$

Condition (*iii*) is thus seen to be tantamount to

$$(18)\qquad \frac{\partial(x, y)}{\partial(t, u)}\neq 0, \qquad t=0, \quad u=0.$$

It follows, then, from Chap. V, § 12, that the first two of the equations (10) can be solved for t and u in terms of x and y, and that the first three equations (10) consequently represent a surface S:

$$(19)\qquad z=\Psi(x, y),$$

where Ψ is continuous, together with its partial derivatives of the first order. Moreover, S contains both curves, $C_{\mathfrak{A}}$ and D. It remains to show that the function (19) satisfies the given differential equation, (1).

The derivatives of this function are given by the formulas (cf. Chap. V, p. 150, Ex. 31):

$$\Psi_x(x, y)=\frac{\partial(z, y)}{\partial(t, u)}\bigg/\frac{\partial(x, y)}{\partial(t, u)}, \qquad \Psi_y(x, y)=\frac{\partial(x, z)}{\partial(t, u)}\bigg/\frac{\partial(x, y)}{\partial(t, u)}.$$

Consider, on the other hand, the equations

$$(20)\qquad \begin{cases} \dfrac{\partial z}{\partial t} = p\dfrac{\partial x}{\partial t} + q\dfrac{\partial y}{\partial t} \\[2ex] \dfrac{\partial z}{\partial u} = p\dfrac{\partial x}{\partial u} + q\dfrac{\partial y}{\partial u}. \end{cases}$$

If these be true, then

$$\Psi_x = p, \qquad \Psi_y = q,$$

and the proof is complete. We proceed, therefore, to establish equations (20).

The first of the equations (20) is true by (16). To prove the second, let

$$U(t, u) = \frac{\partial z}{\partial u} - p\frac{\partial x}{\partial u} - q\frac{\partial y}{\partial u}.$$

When $t = 0$, $U = 0$ by (15), (*ii*). We wish to show that U is 0 for all values of (t, u). This can be shown as follows. We have:

$$\frac{\partial U}{\partial t} = \frac{\partial^2 z}{\partial t\,\partial u} - p\frac{\partial^2 x}{\partial t\,\partial u} - q\frac{\partial^2 y}{\partial t\,\partial u} - \frac{\partial p}{\partial t}\frac{\partial x}{\partial u} - \frac{\partial q}{\partial t}\frac{\partial y}{\partial u}.$$

Differentiating the first equation (20) with respect to u, we have:

$$0 = \frac{\partial^2 z}{\partial u\,\partial t} - p\frac{\partial^2 x}{\partial u\,\partial t} - q\frac{\partial^2 y}{\partial u\,\partial t} - \frac{\partial p}{\partial u}\frac{\partial x}{\partial t} - \frac{\partial q}{\partial u}\frac{\partial y}{\partial t}.$$

On subtracting this equation from its predecessor and reducing by means of (16), *i.e.* (9), we have:

$$\frac{\partial U}{\partial t} = (X + pZ)\frac{\partial x}{\partial u} + (Y + qZ)\frac{\partial y}{\partial u} + P\frac{\partial p}{\partial u} + Q\frac{\partial q}{\partial u}.$$

Finally, differentiate (1) with respect to u:

$$0 = \frac{\partial F}{\partial u} = X\frac{\partial x}{\partial u} + Y\frac{\partial y}{\partial u} + Z\frac{\partial z}{\partial u} + P\frac{\partial p}{\partial u} + Q\frac{\partial q}{\partial u}.$$

On subtracting this equation from the preceding one and reducing, we find:

$$\frac{\partial U}{\partial t} = -ZU.$$

It follows that, if we give to u an arbitrary value and hold it fast, and if we now integrate with respect to t,

$$U = U_0 e^{-\int_0^t Z\,dt}.$$

But $U_0 = U(0, u) = 0$. Hence U is identically 0, q. e. d.

Remark. The usual treatments of this subject follow the historic order of development and begin with the "complete integral" and the "general integral," arriving late (if at all) at the characteristics. Valuable as the historical order is in most subjects in mathematics, the present one is a distinct exception, for the "complete integral" and the "general integral" are artificial, and do not afford an easy or a natural approach to the subject. The *characteristic strips* are the key to the situation, and by means of them the "complete integral" and the "general integral" can be best explained; cf. Goursat, *Cours d'analyse*, vol. II, 2d ed., 1911, p. 593, and Goursat-Bourlet, *Leçons sur l'intégration des équations aux dérivées partielles du premier ordre.*

25. Extension to the Case of $n+1$ Variables. The foregoing treatment admits extension to the case of partial differential equations of the first order with n independent variables:

$$(21)\qquad F(x_1, \cdots, x_n, z, p_1, \cdots, p_n) = 0, \qquad p_k = \frac{\partial z}{\partial x_k};$$

cf. Goursat, cited in § 24. The analogue of the curve D is a manifold of $n-1$ dimensions:

$$D: \qquad z = \phi(u_1, \cdots, u_{n-1}), \quad x_k = \phi_k(u_1, \cdots, u_{n-1}), \qquad k = 1, \cdots, n,$$

which shall not be tangent at A to the curve $C_{\mathfrak{A}}$:

$$\begin{vmatrix} P_1 & \cdots & P_n \\ \dfrac{\partial \phi_1}{\partial u_1} & \cdots & \dfrac{\partial \phi_n}{\partial u_1} \\ \cdot & \cdots & \cdot \\ \dfrac{\partial \phi_1}{\partial u_{n-1}} & \cdots & \dfrac{\partial \phi_n}{\partial u_{n-1}} \end{vmatrix} \neq 0.$$

The functions $p_1, \cdots, p_n$ are determined along this manifold, and substituted, together with the ϕ_k, in the equations which correspond to (10). The requirements (15), (*ii*) and (*iii*), now become:

$$(ii)\quad \frac{\partial z}{\partial u_k} = p_1 \frac{\partial x_1}{\partial u_k} + p_2 \frac{\partial x_2}{\partial u_k} + \cdots + p_n \frac{\partial x_n}{\partial u_k}, \qquad t = 0,\ u = u,\ k = 1, \cdots, n-1.$$

$$(iii)\qquad \begin{vmatrix} P_1 & \cdots & P_n \\ \dfrac{\partial x_1}{\partial u_1} & \cdots & \dfrac{\partial x_n}{\partial u_1} \\ \cdot & \cdots & \cdot \\ \dfrac{\partial x_1}{\partial u_{n-1}} & \cdots & \dfrac{\partial x_n}{\partial u_{n-1}} \end{vmatrix} \neq 0.$$

It is then shown that the function

$$U_k(t, u_1, \cdots, u_{n-1}) = \frac{\partial z}{\partial u_k} - p_1 \frac{\partial x_1}{\partial u_k} - \cdots - p_n \frac{\partial x_n}{\partial u_k},$$

where k is any one of the numbers $1, \cdots, n-1$, satisfies the relation

$$\frac{\partial U_k}{\partial t} = - Z U_k.$$

Hence
$$U_k = U_k^0 e^{-\int_0^t Z\,dt},$$

where $(u_1, \cdots, u_{n-1})$ is arbitrary, but fixed with reference to the integration. But $U_k^0 = 0$ by hypothesis. Hence U_k vanishes identically.

Thus is proved the theorem that there exists one and only one solution of the given differential equation of the form

$$z = \Psi(x_1, \cdots, x_n),$$

where Ψ, together with its partial derivatives of the first order, is continuous throughout the complete neighborhood of the point $(a_1, \cdots, a_n)$, and where the manifold here represented contains both C_M and D.

26. The Equations of Dynamics. It is the theory of the solution of the partial differential equation (1) or (21) by means of the characteristics and the so-called "complete integral" that forms the foundation of the treatment of the motion of a material system with n degrees of freedom according to the methods of Hamilton and Jacobi. The completion of the mathematical theory is found in Goursat, l. c. (§ 24 above). The dynamical problem is discussed in Appell, *Mécanique rationelle*, vol. I, 2d ed., p. 550, and vol. II, 2d ed., p. 407. These methods have recently again come into prominence through their use by the physicists in the study of the atom; cf. Sommerfeld, *Atomic Structure and Spectral Lines.*

27. The Partial Differential Equations of Mathematical Physics. We have met Laplace's Equation:

(*i*)
$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0,$$

in the attraction of gravitating matter and in an electric or magnetic field of force, u denoting the potential function; again, in the irrotational flow of an incompressible fluid, u denoting the velocity potential; and still again in the steady flow of heat or electricity in a homogeneous, isotropic conductor, — the general equation for any (unsteady) flow being

$$\text{(ii)} \qquad \frac{\partial u}{\partial t} = a^2 \left\{ \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} \right\},$$

where u denotes the temperature or the potential.

The equation of the vibrating string is found to be (Chap. XV, § 7 and Chap. XVII, § 11):

$$\text{(iii)} \qquad \frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2},$$

where u denotes the distance of any point of the string from its position of rest, the motion being either transverse or longitudinal

The equation of the vibrating membrane,

$$\text{(iv)} \qquad \frac{\partial^2 u}{\partial t^2} = a^2 \left\{ \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right\},$$

u denoting the transverse displacement, and the sound equation,

$$\text{(v)} \qquad \frac{\partial^2 u}{\partial t^2} = a^2 \left\{ \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} \right\},$$

where u denotes the velocity potential of the vibrating medium, are deduced by Hamilton's Principle, Chap. XVII, § 11.

The foregoing are all linear partial differential equations of the second order, and they, with others like them,* form a set of equations known as the Partial Differential Equations of Mathematical Physics. They are treated by various methods, notably by development into series, by definite integrals, and by integral equations, and their discussion forms an extended theory. We take this occasion to give an example of the first method, since it affords a natural approach to Fourier's Series.

Consider the problem in the flow of heat, formulated in Chap. XI, § 16, as the typical boundary value problem for Laplace's Equation in two dimensions. This problem calls analytically for a function, u, continuous within and on the boundary of the circle

$$x^2 + y^2 = 1,$$

having continuous partial derivatives of the first and second orders which satisfy Laplace's equation

*The partial differential equations of the vibrating rod and the vibrating plate (Chap. XVII, § 11) are also linear, but of higher order.

(1) $$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=0, \quad \text{or} \quad \frac{\partial^2 u}{\partial r^2}+\frac{1}{r}\frac{\partial u}{\partial r}+\frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2}=0,$$

within the circle, and taking on an arbitrarily prescribed set of continuous boundary values, $f(\theta)$, on the circumference, C:

(2) $$u|_C = f(\theta).$$

The following plan of attack is important in applied mathematics. Seek first special solutions of Laplace's equation, written in polar coordinates, which shall be of the form:

(3) $$u = R\Theta,$$

where R is a function of r alone, and Θ a function of θ alone. Substituting in (1) we have:

$$\Theta\left[\frac{d^2R}{dr^2}+\frac{1}{r}\frac{dR}{dr}\right]+\frac{R}{r^2}\frac{d^2\Theta}{d\theta^2}=0, \qquad \text{or} \qquad \frac{\dfrac{d^2R}{dr^2}+\dfrac{1}{r}\dfrac{dR}{dr}}{\dfrac{R}{r^2}}=-\frac{\dfrac{d^2\Theta}{d\theta^2}}{\Theta}.$$

Since one side of this equation is independent of θ, and the other side, independent of r, it follows that each side must be constant; set this constant $= n^2$. We are thus led to the two differential equations:

(4) $$\frac{d^2\Theta}{d\theta^2}+n^2\Theta=0, \qquad r^2\frac{d^2R}{dr^2}+r\frac{dR}{dr}-n^2R=0.$$

The solutions of these differential equations are:

(5) $$\Theta = a\cos n\theta + b\sin n\theta, \qquad \begin{cases} R = Ar^n + Br^{-n}, & 0 < n. \\ R = A + B\log r, & n = 0. \end{cases}$$

Since the function we wish to represent is continuous at the centre of the circle, we should look askance at a solution (3) which became infinite there, and so we set $B = 0$.

Moreover, Θ must be periodic, for its value is the same when θ is increased by 2π. Hence n must be an integer. This property of the periodicity of Θ also justifies our choice of the above constant as positive, $= n^2$. If we had taken it as negative, $= -n^2$, the corresponding differential equation in Θ would not have yielded any periodic solutions ($\not\equiv 0$).

Let n run through the values $0, 1, 2, \cdots$. Then the sum

(6) $$\sum r^n(a_n\cos n\theta + b_n\sin n\theta)$$

is a solution of (1), and the infinite series will be, too, if it converges properly. Suppose it does. Then, along C, we shall have $r=1$, and from (2) and (6)

$$(7) \qquad f(\theta)=\sum_{n=0}^{\infty}(a_n\cos n\theta+b_n\sin n\theta).$$

This last equation suggests the plan of attack. Begin by developing the function $f(\theta)$ into a Fourier's series (Chap. XVI, § 1), thus determining the coefficients a_n and b_n in (7). Next, multiply the general term of (7) by r^n. The series thus obtained will be the solution of the problem:

$$(8) \qquad u=\sum_{n=0}^{\infty}r^n(a_n\cos n\theta+b_n\sin n\theta),$$

provided the series converges properly; for (a) each term in the series satisfies the partial differential equation (1) throughout the interior of the circle; and (b) the limiting function u takes on the prescribed boundary values when $r=1$.

The questions of convergence here are not simple, and their study forms a large and important chapter in modern analysis. The results show that in all cases in which the conditions imposed on the problem are such as are of interest in physics, the series do converge and thus the physicist may apply the method with confidence that it will yield correct results. Cf. Byerly, *Fourier's Series and Spherical Harmonics*, and the recent work of Hilbert and Courant, *Methoden der mathematischen Physik.*

EXERCISE

The differential equation of the vibrating membrane is

$$\frac{\partial^2 u}{\partial t^2}=a^2\left\{\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}\right\}.$$

Consider a drum-head vibrating so that points initially equidistant from the axis always lie on a circle whose plane is perpendicular to the axis and whose centre is in the axis. On introducing polar coordinates we see that $\partial u/\partial\theta=0$, and hence

$$\frac{\partial^2 u}{\partial t^2}=a^2\left\{\frac{\partial^2 u}{\partial r^2}+\frac{1}{r}\frac{\partial u}{\partial r}\right\}.$$

Apply to this equation the method set forth above, letting

$$u=TR.$$

Show that an infinite sequence of particular solutions is obtained if

$$\frac{d^2T}{dt^2} + \lambda^2 a^2 T = 0, \qquad \frac{d^2R}{dr^2} + \frac{1}{r}\frac{dR}{dr} + \lambda^2 R = 0,$$

$$TR = (a_k \cos \lambda_k at + b_k \sin \lambda_k at) J_0(\lambda_k r).$$

The given initial conditions are:

$$u\big|_{t=0} = f(r), \qquad \frac{\partial u}{\partial t}\bigg|_{t=0} = \phi(r),$$

where $f(r)$ and $\phi(r)$ are continuous, and

$$f(h) = 0, \qquad \phi(h) = 0,$$

h being the radius of the drum-head. For convenience, set $h = 1$.

Now, develop $f(r)$ and $\phi(r)$ as follows:

$$f(r) = \sum_{k=0}^{\infty} a_k J_0(\lambda_k r), \qquad \phi(r) = \sum_{k=0}^{\infty} b_k \lambda_k a J_0(\lambda_k r),$$

where $\lambda_0, \lambda_1, \cdots$ are the positive roots of the function $J_0(x)$. Thus the series

$$u = \sum_{k=0}^{\infty} (a_k \cos \lambda_k at + b_k \sin \lambda_k at) J_0(\lambda_k r)$$

will give the desired solution provided it converges suitably.

Explain the reason for choosing the constant here as λ^2, and not $-\lambda^2$.

CHAPTER XV

ELASTIC VIBRATIONS

1. Simple Harmonic Motion. The simplest case of oscillatory motion about a position of equilibrium is that of *Simple Harmonic Motion*, studied in detail in the chapter on Mechanics, *Introduction to the Calculus*, p. 364; cf. in particular Ex. 7, p. 368.

This case is typical for the great majority of systems with one degree of freedom (one coordinate) such as one meets in physics, so far as a first approximation is concerned. Even continuous media, like a vibrating piano string, can obey this law, and a study of their motion in this simplest case is often a convenient approach to their theory.

The differential equation which dominates simple harmonic motion is

$$\frac{d^2x}{dt^2} + n^2x = 0. \tag{1}$$

We solved it at the time by a method which was useful and suggestive at that stage. The method best adapted for the study of the problems of this chapter is the one set forth in Chap. XIV, § 12. Setting

$$x = e^{mt}$$

we find, for the determination of m, the equation:

$$m^2 + n^2 = 0,$$

and hence the general solution of (1) can be written in either of the forms:

$$x = A\cos nt + B\sin nt \qquad \text{or} \qquad x = C\cos(nt+\gamma) \tag{2}$$

$$[x = C\sin(nt+\gamma)].$$

The Period. The period of a half-oscillation is seen to be π/n; the period from phase to phase is

$$T = \frac{2\pi}{n}. \tag{3}$$

The *amplitude* is constant and $= 2C$.

A second approximation is introduced when we take *damping* into account. This dissipative force is due to the resistance of the atmosphere, or the viscosity of the substance, or other similar causes. It is studied in § 2.

Finally comes the case of *forced vibrations*, studied in § 4 in its simplest form.

EXERCISE

Show that the velocity with which the particle passes through the point of equilibrium is proportional to the amplitude, and compute the kinetic energy which it has at that point.

2. Damping. The physical picture which it is convenient to use in these paragraphs is that of an elastic wire, or spring, its upper end fastened at a point, A, and a weight m attached to its lower end. Then

$$(1) \qquad T = \lambda \frac{s}{l} \qquad\qquad mg = \lambda \frac{s_0}{l},$$

where $l = \overline{AB}$ is the natural length of the string, O is the point of equilibrium, $\overline{BP} = s$, and $\overline{BO} = s_0$. Let $\overline{OP} = x$, or

$$(2) \qquad x = s - s_0.$$

The damping is a force which acts in the direction opposite to that of the motion, and which increases with the velocity. The simplest mathematical formula which will yield such a force is

$$-k\frac{ds}{dt},$$

where k is a small positive constant. And now it turns out physically that, in the case of small oscillations about the position of equilibrium, this formula gives a satisfactory approximation.

From Newton's Second Law of Motion, *Introduction to the Calculus*, p. 348:

$$(3) \qquad m\frac{d^2s}{dt^2} = mg - T - k\frac{ds}{dt}.$$

From (1) and (2) we have

$$T - mg = \lambda\frac{s}{l} - \lambda\frac{s_0}{l} = \frac{\lambda}{l}x,$$

$$\frac{ds}{dt} = \frac{dx}{dt}, \qquad \frac{d^2s}{dt^2} = \frac{d^2x}{dt^2}.$$

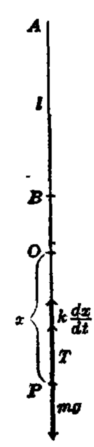


FIG. 88

Hence (3) can be written:

$$m\frac{d^2x}{dt^2}=-\frac{\lambda}{l}x-k\frac{dx}{dt},$$

or

(4) $$\frac{d^2x}{dt^2}+\kappa\frac{dx}{dt}+n^2x=0,$$

where $\kappa=k/m$ and $n=\sqrt{\lambda/ml}.$

Thus we have the differential equation of simple harmonic motion, § 1, (1), to which the damping term, $\kappa\, dx/dt$, has been added.

To integrate equation (4) let $x=e^{mt}$; Chap. XIV, § 12. The equation for m becomes:

$$m^2+\kappa m+n^2=0.$$

Since κ is small, the roots are imaginary. Let

$$\nu=\sqrt{n^2-\tfrac{1}{4}\kappa^2}.$$

The general solution of (4) can now be written in the form:

(5) $$x=Ce^{-\frac{\kappa}{2}t}\cos(\nu t+\gamma),$$

where C and γ are the constants of integration.

EXERCISE

Show that, if the particle is started from a point at which $x=a$ with a velocity equal to u when $t=0$,

$$\tan\gamma=-\frac{1}{\nu}\left(\frac{u}{a}+\frac{\kappa}{2}\right).$$

Choosing γ so that $0\leq\gamma<\pi$, find C.

3. Discussion of the Result. From the solution (5) of § 2 we see that, no matter how the system be set in motion, the particle passes periodically through the point O of no force, the period being π/ν. But the period from phase to phase is twice as great, or*

(6) $$T=\frac{2\pi}{\nu}.$$

Since κ is small, we have (cf. *Introduction to the Calculus*, Infinite Series, p. 413, (2)):

$$\frac{1}{\nu}=\frac{1}{n}\left(1-\frac{\kappa^2}{4n^2}\right)^{-\frac{1}{2}}=\frac{1}{n}\left(1+\frac{\kappa^2}{8n^2}+\cdots\right).$$

* We are here using the letter T for the *time* corresponding to a complete period.

Hence the period (6) differs from the period the same system has when there is no damping, § 1, (3), by a small quantity of the *second* order, referred to κ/n as of the first order:

$$\frac{2\pi}{\nu} = \frac{2\pi}{n} + \left(\frac{\kappa}{n}\right)^2 \left\{ \frac{\pi}{4n} + \text{a small quantity} \right\}.$$

The amplitude of the oscillation dies down, owing to the exponential factor, and approaches 0 as its limit.

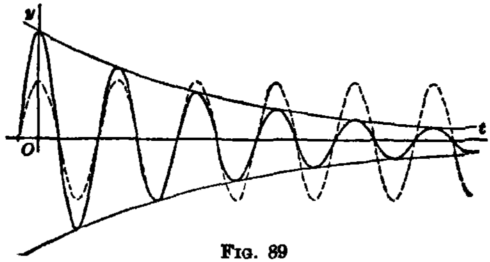


Fig. 89

4. Forced Vibrations. The phenomenon of forced vibrations is familiar to the race through varied manifestations. A regiment of soldiers, in crossing a bridge, is commanded to break step. The chances are that it is unnecessary to do so. But if the natural note, or period, of the bridge should be about the same as the beat of their steps, serious consequences might ensue, for the bridge could be brought into violent vibration.

We are told, too, how the piper fiddled down the bridge by striking the note of the cables, and the walls of Jericho are reported to have fallen in a similar manner.* We have all had the experience of sneezing in a room where there was a banjo, and then hearing the banjo sneeze, too.

The tides form another example, for they are due to the attraction of the sun and the moon.

One of the cheerful recollections of my school days is that of shaking the room in the old Rice Grammar School in Boston. A child, sitting at his desk, with the ball of the foot on the floor could, by causing the leg to move up and down with a period nearly equal to the natural period of the floor, produce vibrations most disturbing to the lady school teacher.

* Joshua vi. 20. It was Mr. Fulton Cutting who called my attention to this fact years ago in Mathematics 5.

The phenomenon can be studied effectively mathematically by means of the following experiment.

Let a soft spring be attached from below to the weight of § 2, and let the lower end of this spring be driven periodically up and down. Thus a periodic force is impressed on the weight, and the motion which ensues is called a *forced vibration.*

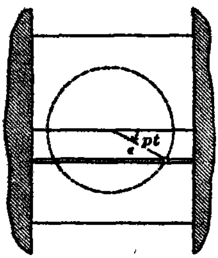


FIG. 90

Imagine a plate which can move vertically and which is provided with a horizontal slot. Behind the plate is a crank which is driven with uniform angular velocity in a plane parallel to that of the plate. The crank carries a pin which passes through the slot, thus causing the plate to move up and down. The depth of the plate below its position when the crank is horizontal to the right ($t=0$) is

$$\epsilon \sin pt.$$

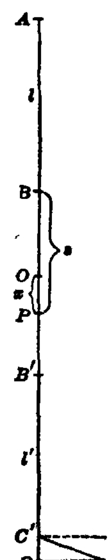


FIG. 91

It is to this plate that the lower end of the soft spring is attached. Let

$$T' = \lambda' \frac{s'}{l'}$$

be the tension in the soft spring. Newton's Second Law becomes:

$$(1) \qquad m\frac{d^2 s}{dt^2} = mg - T + T' - k\frac{ds}{dt}.$$

From the figure we see that the stretched length of the soft spring can be expressed in two ways:

$$l' + s' = \overline{QC'} + \overline{C'B'} + \overline{B'P},$$

where C' denotes the position of the lower end of the soft spring when $t=0$. Now,

$$\overline{QC'} = \epsilon \sin pt, \qquad \overline{C'B'} = l', \qquad \overline{B'P} = \overline{B'O} - x.$$

Moreover, since O is the point of no force,

$$(2) \quad mg + T_0' = T_0 \qquad \text{or} \qquad mg - \frac{\lambda}{l}\overline{BO} + \frac{\lambda'}{l'}\overline{B'O} = 0.$$

Hence the right-hand side of (1) becomes:

$$(3) \qquad mg - \frac{\lambda}{l}(\overline{BO} + x) + \frac{\lambda'}{l'}(\overline{B'O} - x) + \frac{\lambda'}{l'}\epsilon \sin pt - k\frac{dx}{dt}.$$

Thus (1) reduces by the aid of (3) and (2) to

(4) $$m\frac{d^2x}{dt^2} = -\left(\frac{\lambda}{l}+\frac{\lambda'}{l'}\right)x + \frac{\lambda'}{l'}\epsilon \sin pt - k\frac{dx}{dt}$$

or

(5) $$\frac{d^2x}{dt^2} + \kappa\frac{dx}{dt} + n^2x = E \sin pt.$$

5. Integration of the Differential Equation. We can effect the complete integration of the differential equation which governs the motion, § 4, (5), if we find one single special solution; cf. Chap. XIV, § 11. Now, it was long since known or surmised that the system on which a periodic impressed force acts ultimately gives up its own note and takes on the period of the impressed force. . But the *phase* of the one oscillation is different from that of the other; the tides lag behind the moon.

We are thus moved to try an experiment and see if we cannot determine a particular periodic solution of (5), § 4, built on the simplest lines imaginable. So we set

(6) $$x = A \sin (pt - \alpha),$$

when A and α are undetermined constants, and try to determine these so that (6) will be a solution.

Substituting the function (6) in equation (5), we find:

$$A(n^2-p^2)\sin(pt-\alpha) + A\kappa p \cos(pt-\alpha) = E \sin pt.$$

This equation is equivalent to the following:

$$\{A(n^2-p^2)\cos\alpha + A\kappa p \sin\alpha - E\}\sin pt$$
$$-\{A(n^2-p^2)\sin\alpha - A\kappa p\cos\alpha\}\cos pt = 0.$$

The latter equation will be true for all values of t if

(7) $$\begin{cases} A(n^2-p^2)\cos\alpha + A\kappa p\sin\alpha = E \\ A(n^2-p^2)\sin\alpha - A\kappa p\cos\alpha = 0. \end{cases}$$

From the last equation follows that

(8) $$\tan\alpha = \frac{\kappa p}{n^2-p^2}.$$

We will agree to understand by α that root of this equation for which

$$0 < \alpha < \pi.$$

The first equation (7) gives

(9) $$A = \frac{E\kappa p}{\{(n^2 - p^2)^2 + \kappa^2 p^2\} \sin \alpha}.$$

Thus A is always positive.

The experiment has succeeded. We have found a solution of the form (6), where α and A are determined by (8) and (9). The general solution of (5), § 4, can now be written down:

(10) $$x = Ce^{-\frac{\kappa}{2}t} \cos(\nu t + \gamma) + A \sin(pt - \alpha),$$

where C and γ are the constants of integration, and $\nu = \sqrt{n^2 - \frac{1}{4}\kappa^2}$.

6. Discussion of the Result. (*a*) *The system gives up its natural period and takes on the period of the impressed force.*

For, the first term on the right of (10) becomes insignificant, as t increases, no matter how the system was started; the second term, however, is periodic.

(*b*) *When p and ν are nearly, but not quite, equal, beats appear.*

For, there comes an interval in which the arches of the two component curves, (5) of § 2 and (6) of § 5, lie on the same side of the axis of x and have almost coincident bases. Thus they reinforce each other. Then comes, a little later, an interval in which these arches lie on opposite sides of the axis of x and have almost coincident bases. Now, they tend to neutralize each other. And so on. Finally, both phenomena are flattened out as the first term on the right of (10) tends to disappear.

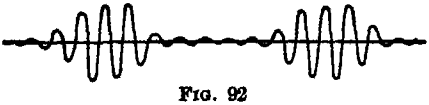


FIG. 92

If the vibration of the system is such as to produce sound, the sound will be loud during the first interval, low during the second, then loud again, and so on, — for the intensity of the sound is greater when the amplitude of the oscillations is greater. Thus we have the phenomenon of beats in acoustics.

(*c*) *When p and ν are equal (or very nearly so) and κ/p is small, the amplitude of the forced vibration is large, and the lag is a quarter-period (nearly).*

Since $\nu^2 = n^2 - \frac{1}{4}\kappa^2$, we have here $(p = \nu)$:

$$n^2 - p^2 = \tfrac{1}{4}\kappa^2, \qquad \tan \alpha = 4\frac{p}{\kappa},$$

and hence α is slightly less than 90°. Moreover,

$$A = \frac{E}{\sin\alpha}\cdot\frac{\kappa p}{\frac{1}{16}\kappa^4+\kappa^2p^2} = \frac{E}{p^2\sin\alpha}\left[1+\tfrac{1}{16}\left(\frac{\kappa}{p}\right)^2\right]^{-1}\Big/\frac{\kappa}{p}$$

or A is nearly inversely proportional to κ/p.

Now, in the system described in § 4, E is small. If, however, E/p^2 is not small in comparison with κ/p, the value of A will be large, and thus the system on which the periodic force is impressed will oscillate violently. This is the case in which the natural period of the bridge is nearly the same as the period of the force impressed on it by the regiment; or the case of the banjo that sneezes sympathetically.

For a further study of the subject of these last paragraphs the student is referred to Lord Rayleigh's *Theory of Sound*, vol. I, Chap. 3, and to Helmholtz, *Theoretische Physik*, vol. 3, pp. 1–71.

7. The Differential Equation of the Vibrating String. Consider a perfectly flexible homogeneous string of uniform density, which obeys Hooke's Law, and which has its two ends fastened at two fixed points, A and B, further apart than the natural length of the string. The string is given an arbitrary initial displacement, subject merely to the conditions that it is nowhere stretched beyond the elastic limit and that it is nowhere slack, and that certain requirements of continuity are observed. To determine the subsequent motion. A plucked violin string, or a piano string, struck by the hammer, suggest the sort of problem that is meant.

Hooke's Law. Let l_0 denote the natural length of the string, and let $l_1 = AB$ (Fig. 93). Then the tension is given by the formula:

(1) $$T = \lambda\frac{l-l_0}{l_0}.$$

In particular, the tension in the string at rest is

(2) $$T_1 = \lambda\frac{l_1-l_0}{l_0}.$$

Let x be the coordinate of any point P of the string when in equilibrium, measured from A, or $x = AP$. Consider a segment $PP' = \Delta x$, and let h be the unstretched length of this segment. Then

(3) $$T_1 = \lambda\frac{\Delta x - h}{h}.$$

Let the end A of the string still be held fast, but let the other end be pulled out to a point C. Let Q, with the coordinate ξ, be the point into which P is carried. The segment PP', of length Δx, goes over into QQ', of length $\Delta\xi$, and since the natural length of QQ' is also h, we have for the new tension

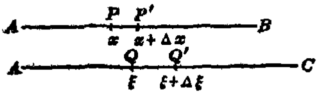


FIG. 93

(4) $$T_2 = \lambda \frac{\Delta\xi - h}{h}.$$

Eliminating h between (3) and (4) gives:

(5) $$T_2 = T_1 \frac{\Delta\xi}{\Delta x} + \lambda\left(\frac{\Delta\xi}{\Delta x} - 1\right).$$

Let Δx approach 0:

(6) $$T_2 = T_1 \frac{d\xi}{dx} + \lambda\left(\frac{d\xi}{dx} - 1\right).$$

Let u denote the distance PQ, or

(7) $$u = \xi - x.$$

Then

(8) $$T_2 = T_1 + (T_1 + \lambda)\frac{du}{dx}.$$

This last equation expresses Hooke's Law for the tension at an arbitrary point P in terms of the constants T_1 and λ and the *rate of stretching*, du/dx, at P. In this form, the law admits extension to the case of variable tension such as arises in a vibrating string; cf. Formulas (16) and § 8, (6), below.

Longitudinal Vibrations. Consider now the particular case of the general problem proposed above, in which the string is displaced along its own line, the ends remaining fixed; and then released. This initial displacement is defined by an equation

(9) $$\xi = f(x), \qquad 0 \leqq x \leqq l_1,$$

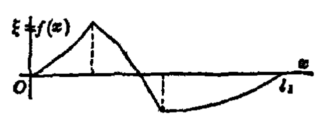


FIG. 94

where f is a continuous function. We will assume, furthermore, that f has a derivative which is also continuous, except possibly at a finite number of points, at each of which a forward derivative and a backward derivative exist.

Our problem is: *To determine the position, at an arbitrary instant t, of that point of the string which has the coordinate x when the string is at rest in equilibrium; i.e.* to determine the functions

$$\xi = \xi(x, t), \qquad u = u(x, t),$$

where u and ξ are connected by (7).

To do this, consider a segment QQ' of the string, and isolate this system. Let PP' be the corresponding segment when the string is in equilibrium. Its mass is $m = \rho\,\Delta x$. Its centre of gravity, $\bar{\xi}$, is given by the formula

$$(10) \qquad \bar{\xi} = \frac{\int_{\xi}^{\xi+\Delta\xi} \rho(\xi)\,\xi\,d\xi}{m},$$

where $\rho(\xi)$ denotes the density at any point ξ of QQ'.

Change the variable of integration from ξ to x. Then

$$\rho(\xi)\,d\xi = \rho\,dx,$$

where ρ is the constant density of the string in equilibrium. For, the mass of the string from A to any point ξ is

$$\int_0^{\xi} \rho(\xi)\,d\xi = \rho x. \qquad \text{Hence} \quad \rho(\xi)\frac{d\xi}{dx} = \rho.$$

Thus (10) becomes:

$$(11) \qquad m\bar{\xi} = \int_x^{x+\Delta x} \rho\,\xi(x, t)\,dx,$$

where t is arbitrary, but constant.

Motion of the Centre of Mass. It is a fundamental law of mechanics that the centre of mass of any material system (system of particles, rigid body, or even the Mississippi River) moves as if all its mass were concentrated there and all the forces were replaced by equal forces acting there.

Applying this principle to the segment QQ' of the string, we have:

$$(12) \qquad m\frac{\partial^2 \bar{\xi}}{\partial t^2} = T' - T.$$

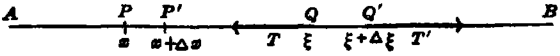


FIG. 95

If, now, we assume that $\xi(x, t)$ is, together with its partial derivatives ξ_t, ξ_{tt}, continuous in the two independent variables x and t, then the integral (11) can be differentiated by Leibniz's Rule, Chap. XIX, § 1, and we have:

(13) $$m\frac{\partial^2\bar{\xi}}{\partial t^2}=\int_x^{x+\Delta x}\rho\frac{\partial^2\xi}{\partial t^2}dx=\rho\,\Delta x\,\xi_{tt}(x+\theta\Delta x,t),\qquad 0<\theta<1.$$

On the other hand, $T=T(x,t)$, and

$$T'-T=T(x+\Delta x,t)-T(x,t)\;=\;\Delta x\,T_x(x+\theta'\Delta x,t),$$

if, as we will assume, $T(x,t)$, together with $\partial T/\partial x$, is continuous in (x,t). Thus (12) becomes:

(14) $$\rho\Delta x\left(\frac{\partial^2\xi}{\partial t^2}\right)_{(x+\theta\Delta x,\,t)}=\Delta x\,T_x(x+\theta'\Delta x,t).$$

Dividing by Δx and allowing Δx to approach 0, we find:

(15) $$\rho\frac{\partial^2\xi}{\partial t^2}=\frac{\partial T}{\partial x},\qquad\text{or}\qquad\rho\frac{\partial^2 u}{\partial t^2}=\frac{\partial T}{\partial x}.$$

As yet, no other hypothesis than that of continuity (including the derivative) has been made regarding T. We now assume Hooke's Law, which here takes the form: *

(16) $$T=T_1+c\,\frac{\partial u}{\partial x},\qquad\qquad c=T_1+\lambda.$$

Thus, finally, we obtain from (15)

(17) $$\frac{\partial^2 u}{\partial t^2}=a^2\frac{\partial^2 u}{\partial x^2},\qquad\qquad a^2=\frac{c}{\rho}.$$

This is the equation of the vibrating string. It is exact, if Hooke's Law is exact, no approximations of any sort having entered. There are, however, certain implicit restrictions, which consist in the assumption that $\partial u/\partial x$ never becomes so large that the elastic limit of the string is surpassed; and also that the string never becomes slack and buckles.

8. Continuation; the General Case. Let the string now be displaced in any curve, suitably restricted with respect to continuity, and let its form be given, at any instant after its release, by the equations:

(1) $$\xi=\xi(x,t),\qquad\qquad \eta=\eta(x,t);$$

the motion to take place in a fixed plane. Let

$$u=\xi-x,\qquad\qquad v=\eta-0.$$

* This last statement is not a mathematical inference, but a new physical law, suggested, it is true, by the mathematical deduction of (8) from (1).

Consider an arc PP' of the string. The external forces acting on it are: T at P and T' at P', each along the tangent. The motion of its centre of gravity is given by the equations:

(2) $$m\frac{\partial^2\bar{\xi}}{\partial t^2} = T'\cos\tau' - T\cos\tau; \qquad m\frac{\partial^2\bar{\eta}}{\partial t^2} = T'\sin\tau' - T\sin\tau.$$

For $\bar{\eta}$ we have:

$$\bar{\eta} = \frac{\int_s^{s+\Delta s} \rho(s)\,\eta(x,t)\,ds}{m}.$$

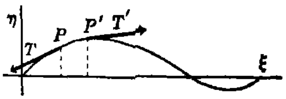


FIG. 96

Here, t is constant, and ds is given by (1):

(3) $$ds^2 = (\xi_x^2 + \eta_x^2)\,dx^2.$$

Moreover, $$\rho(s)\,ds = \rho\,dx.$$
Hence

$$\int_s^{s+\Delta s} \rho(s)\,\eta\,ds = \int_x^{x+\Delta x} \rho\,\eta\,dx,$$

(4) $$m\frac{\partial^2\bar{\eta}}{\partial t^2} = \int_x^{x+\Delta x} \rho\frac{\partial^2\eta}{\partial t^2}\,dx = \Delta x\cdot\rho\left.\frac{\partial^2\eta}{\partial t^2}\right|_{(x+\theta\Delta x,\,t)}.$$

On substituting this value in (2), dividing by Δx, and taking limits, we have:

(5) $$\rho\frac{\partial^2\eta}{\partial t^2} = \frac{\partial}{\partial x}(T\sin\tau).$$

Hooke's Law here takes on the form, cf. (6), § 7:

(6) $$T = (T_1+\lambda)\frac{\partial s}{\partial x} - \lambda = c\sqrt{\xi_x^2+\eta_x^2} - \lambda, \qquad c = T_1 + \lambda.$$

Thus $$T\sin\tau = c\eta_x - \frac{\lambda\eta_x}{\sqrt{\xi_x^2+\eta_x^2}}.$$

On introducing u and v we have:

(7) $$\rho\frac{\partial^2 v}{\partial t^2} = \frac{\partial}{\partial x}\left[cv_x - \frac{\lambda v_x}{\sqrt{(1+u_x)^2+v_x^2}}\right].$$

The corresponding equation for $\partial^2 u/\partial t^2$ is:

(8) $$\rho\frac{\partial^2 u}{\partial t^2} = \frac{\partial}{\partial x}\left[cu_x - \frac{\lambda(1+u_x)}{\sqrt{(1+u_x)^2+v_x^2}}\right].$$

Approximations. Equation (7) is ordinarily replaced by an approximate equation in case u_x and v_x (and hence also u) are numerically small.

$$\frac{\partial}{\partial x}\frac{v_x}{\sqrt{(1+u_x)^2+v_x^2}}=\frac{v_{xx}}{\sqrt{(1+u_x)^2+v_x^2}}-\frac{[(1+u_x)u_{xx}+v_xv_{xx}]\,v_x}{[(1+u_x)^2+v_x^2]^{\frac{3}{2}}}.$$

Dropping terms of the order of magnitude of u_x and v_x we have:

(9) $$\frac{\partial^2 v}{\partial t^2}=a^2\frac{\partial^2 v}{\partial x^2},\qquad a^2=\frac{T_1}{\rho}.$$

Equation (8) becomes

(10) $$\frac{\partial^2 u}{\partial t^2}=b^2\frac{\partial^2 u}{\partial x^2},\qquad b^2=\frac{T_1+\lambda}{\rho}.$$

EXERCISES

1. Show that, when the string vibrates in three dimensions and we set

$$u=\xi-x,\qquad v=\eta,\qquad w=\zeta,$$

equation (7) becomes

(11) $$\rho\,\frac{\partial^2 v}{\partial t^2}=\frac{\partial}{\partial x}\left[c\,v_x-\frac{\lambda v_x}{\sqrt{(1+u_x)^2+v_x^2+w_x^2}}\right].$$

Write down the other two equations.

2. Show that equations (7) and (8) hold when ρ is any continuous function of x, and also when the end B is not fixed, provided that no external forces act along the string.

3. Show that, if there is a slight damping, which is nearly proportional to the length of a short arc and to the component of its velocity perpendicular to the axis of x, and is nearly perpendicular to this axis, the approximate equation for small vibrations becomes:

(12) $$\frac{\partial^2 v}{\partial t^2}+\kappa\frac{\partial v}{\partial t}=a^2\frac{\partial^2 v}{\partial x^2}.$$

4. An inextensible heavy chain hangs from a fixed point, A, and is displaced in a vertical plane. If gravity and the force at A are the only forces which act, show that the motion is governed by the differential equations:

$$\rho\frac{\partial^2 u}{\partial t^2}=\frac{\partial}{\partial x}\left[T\left(1+\frac{\partial u}{\partial x}\right)\right]+g\rho,\qquad \rho\frac{\partial^2 v}{\partial t^2}=\frac{\partial}{\partial x}\left(T\frac{\partial v}{\partial x}\right),$$

$$\left(1+\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial v}{\partial x}\right)^2=1.$$

5. For slight displacements, show that the approximate equations in Question 4 are:

$$T = g\rho(l - x), \qquad \frac{\partial^2 v}{\partial t^2} = g\frac{\partial}{\partial x}\left[(l - x)\frac{\partial v}{\partial x}\right];$$

or, on setting $l - x = x'$, and then dropping the accent:

$$\frac{\partial^2 v}{\partial t^2} = g\frac{\partial}{\partial x}\left(x\frac{\partial v}{\partial x}\right).$$

6. A heavy uniform string of natural length l, obeying Hooke's Law, hangs from one end, at rest under gravity. Show that it is stretched by the length $g\rho l^2/(2\lambda)$.

7. If the string of Question 6 is flexible, and if it vibrates in a vertical plane, find the differential equations which govern the motion.

9. The Differential Equation of the Vibrating Membrane. Let an elastic membrane, like a drum head, be clamped along a plane curve. Let it be displaced and released. To determine the differential equation of the motion.

Hitherto we have derived the differential equation without making approximations, and we have proceeded from it to the approximate differential equation. In the present case, this course would involve too extended a treatment of the theory of elasticity. We will confine ourselves to the simplest case, assuming that the motion of each point of the membrane is orthogonal to the plane of the bounding curve, C; that the displacement, u, of the points of the membrane, together with $\partial u/\partial x$ and $\partial u/\partial y$, is small, the (x, y)-plane coinciding with that of C; and finally, that the tension, T, is the same in all directions at any given point, and is constant at all points.

Consider a piece, Σ, of the membrane, whose projection on the (x, y)-plane is S. For the motion of its centre of mass we have approximately:

$$(1) \qquad m\frac{\partial^2 \bar{u}}{\partial t^2} = \int_\Gamma T\frac{\partial u}{\partial n}\,ds,$$

where n is the outer normal to Γ in the (x, y)-plane.*

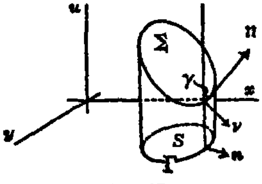


FIG. 97

* We cannot properly speak of the "derivative along the outer normal" in this case. What we mean by this expression is the *negative* of the derivative along the inner normal.

By Chap. XI, § 4:

$$\int_{\Gamma}\frac{\partial u}{\partial n}\,ds=\iint_S\left(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}\right)dS.$$

On the other hand,

$$m\bar{u}=\iint_S \rho\, u\, dS.$$

Hence (1) becomes:

(2) $$\iint_S\left[\rho\frac{\partial^2 u}{\partial t^2}-T\left(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}\right)\right]dS=0.$$

Since S is arbitrary, it follows that

(3) $$\frac{\partial^2 u}{\partial t^2}=a^2\left(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}\right),\qquad a^2=\frac{T}{\rho},$$

and this is the differential equation of the vibrating membrane.

EXERCISES

1. If there is damping, represented by a force per unit of area proportional to $\partial u/\partial t$ and at right angles to the plane of C, show that the differential equation will be:

$$\frac{\partial^2 u}{\partial t^2}+\kappa\frac{\partial u}{\partial t}=a^2\left(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}\right),\qquad a^2=\frac{T}{\rho}.$$

2. If the membrane is heavy and flexible, and the axis of u is directed upward, show that

$$\frac{\partial^2 u}{\partial t^2}=a^2\left(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}\right)-g.$$

3. Consider the case of variable tension, T, but assume that, at any point, the membrane "pulls equally in all directions." Define accurately the *specific tension*.

4. Write down the accurate integral which corresponds to the approximate integral on the right of equation (1). Hence derive the approximate integral.

CHAPTER XVI

FOURIER'S SERIES AND ORTHOGONAL FUNCTIONS

1. Formal Development into a Fourier's Series. By a *Fourier's Series* is meant a series of the form:

(1) $$\frac{a_0}{2}+\sum_{n=1}^{\infty}(a_n\cos nx+b_n\sin nx).$$

Its terms all admit the period 2π, and so it is sufficient to consider the series in an interval of length 2π. The simplest such intervals are

(2) $$0\leqq x<2\pi \qquad\text{and}\qquad -\pi<x\leqq\pi.$$

Let $f(x)$ be a function which is continuous in the second of these intervals, or has at most a finite number of finite discontinuities as shown in the figure.* It is convenient now to extend the definition of the function to all values of x by the property of periodicity:

$$f(x+2\pi)=f(x).$$

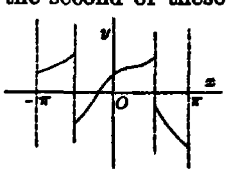


FIG. 98

It is a theorem, the proof of which cannot be taken up here, that such a function can be represented by a Fourier's series:

(3) $$f(x)=\tfrac{1}{2}a_0+a_1\cos x+b_1\sin x+a_2\cos 2x+b_2\sin 2x$$
$$+a_3\cos 3x+b_3\sin 3x+\cdots,$$

i.e. this series, when the coefficients are properly determined, will converge for every value of x, and its value will coincide with that of the function at the point.

* More precisely, it shall be possible to divide the interval into a finite number of segments by the points $x_0=-\pi,\ x_1,\ \cdots,\ x_{n-1},\ x_n=\pi$ such that, in the interval $x_{k-1}<x<x_k$ the function and its first derivative are finite and continuous, and each approaches a limit at either end of the interval. Finally, at a point of discontinuity, the function shall be given the arithmetic mean of the two limiting values which it approaches from either side; and its value for $x=\pi$ shall likewise be the mean of the two limiting values when x approaches $-\pi$ or π.

It is further shown that, when the above series is multiplied by $\cos nx$ or $\sin nx$, the new series can be integrated term-by-term (*i.e.* just as if it were a finite sum) throughout the interval $(-\pi, \pi)$.

This last property enables us to compute the coefficients. Observe that

$$(4) \qquad \int_{-\pi}^{\pi} \sin mx \cos nx\, dx = 0, \qquad \int_{-\pi}^{\pi} \sin mx \sin nx\, dx = 0,$$

$$\int_{-\pi}^{\pi} \cos mx \cos nx\, dx = 0,$$

where m, n are any two integers, positive or zero, in the first relation, but in the second and third, $m \neq n$. For,

$$\sin mx \cos nx = \tfrac{1}{2}[\sin (m-n)x + \sin (m+n)x],$$

$$\sin mx \sin nx = \tfrac{1}{2}[\cos (m-n)x - \cos (m+n)x],$$

$$\cos mx \cos nx = \tfrac{1}{2}[\cos (m-n)x + \cos (m+n)x].$$

Moreover,

$$(5) \qquad \int_{-\pi}^{\pi} \sin^2 nx\, dx = \pi, \qquad \int_{-\pi}^{\pi} \cos^2 nx\, dx = \pi, \qquad \int_{-\pi}^{\pi} \tfrac{1}{2}\, dx = \pi.$$

On multiplying equation (3) through by $\cos nx$, $n = 0, 1, 2, \cdots$, and integrating the resulting equation from $-\pi$ to π, all the terms but one on the right drop out by virtue of (4), and we have:

$$\int_{-\pi}^{\pi} f(x) \cos nx\, dx = a_n \int_{-\pi}^{\pi} \cos^2 nx\, dx = \pi a_n.$$

Similarly, on multiplying through by $\sin nx$ and integrating, we have:

$$\int_{-\pi}^{\pi} f(x) \sin nx\, dx = b_n \int_{-\pi}^{\pi} \sin^2 nx\, dx = \pi b_n.$$

Thus the coefficients are determined, and we have:

$$(6) \qquad \begin{cases} a_n = \dfrac{1}{\pi} \displaystyle\int_{-\pi}^{\pi} f(x) \cos nx\, dx, & n = 0, 1, 2, \cdots; \\[2ex] b_n = \dfrac{1}{\pi} \displaystyle\int_{-\pi}^{\pi} f(x) \sin nx\, dx, & n = 1, 2, \cdots. \end{cases}$$

The series (3), in which the coefficients a_n and b_n have the values given by (6), is known as the *formal development* of the function $f(x)$, for this series exists in form, quite apart from the question of whether it converges and represents the function.

Example 1. Let $f(x) = x, \quad -\pi < x < \pi; \quad f(\pi) = 0.$ Then

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} x\,dx = 0;$$

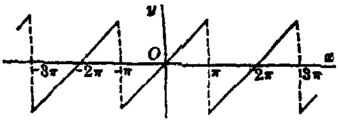


FIG. 99

and *Tables*, Nos. 336, 340:

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} x\cos nx\,dx = \frac{1}{\pi n^2}\left(\cos nx + nx\sin nx\right)_{-\pi}^{\pi} = 0, \qquad n = 1, 2, \cdots;$$

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} x\sin nx\,dx = \frac{1}{\pi n^2}\left(\sin nx - nx\cos nx\right)_{-\pi}^{\pi} = -\frac{2\cos n\pi}{n}.$$

Hence

$$x = 2\left[\frac{\sin x}{1} - \frac{\sin 2x}{2} + \frac{\sin 3x}{3} - \cdots\right]. \tag{7}$$

Example 2. Let $f(x) = x, \quad 0 < x \leqq \pi; \; f(x) = -x, \quad -\pi < x \leqq 0.$

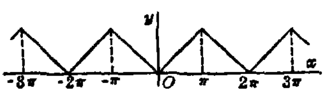


FIG. 100

Here,

$$a_0 = \pi; \qquad a_n = \frac{2\cos n\pi - 2}{\pi n^2}; \qquad b_n = 0$$

Hence

$$f(x) = \frac{\pi}{2} - \frac{4}{\pi}\left[\frac{\cos x}{1^2} + \frac{\cos 3x}{3^2} + \frac{\cos 5x}{5^2} + \cdots\right].$$

EXERCISES *

Obtain the formal development into a Fourier's series in each of the following cases, and plot the curve.

1. $f(x) = 1, \qquad 0 < x < \pi;$
   $f(x) = -1, \quad -\pi < x < 0;$
   $f(0) = f(\pi) = 0.$

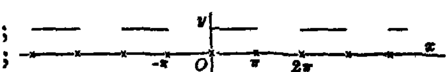


FIG. 101

$$\textit{Ans.}\ f(x) = \frac{4}{\pi}\left[\frac{\sin x}{1} + \frac{\sin 3x}{3} + \frac{\sin 5x}{5} + \cdots\right].$$

* For further examples, cf. Byerly, *Fourier's Series*, pp. 41–51.

2. $f(x)=x, \quad 0<x<\pi; \qquad f(x)=0, \quad -\pi<x\leqq 0; \quad f(\pi)=\frac{1}{2}\pi.$

3. $$f(x)=x^2, \quad -\pi<x\leqq\pi.$$

4. $f(x)=x^2, \quad 0<x<\pi; \qquad f(x)=-x^2, \quad -\pi<x\leqq 0;$
$$f(\pi)=0.$$

5. Prove that, if $f(x)$ is an odd function in the interval $-\pi<x<\pi$; *i.e.* if $f(-x)=-f(x)$ (and if $f(\pi)=0$), then $a_n=0$, $n=0, 1, 2, \cdots$, and

$$f(x)=b_1\sin x+b_2\sin 2x+b_3\sin 3x+\cdots.$$

But if $f(x)$ is an even function in that interval, *i.e.* if $f(-x)=f(x)$, then $b_n=0$, $n=1, 2, \cdots$, and

$$f(x)=\tfrac{1}{2}a_0+a_1\cos x+a_2\cos 2x+a_3\cos 3x+\cdots.$$

6. Show that a function $f(x)$ which, in the interval $0\leqq x\leqq\pi$, satisfies the conditions of the text, can be developed formally (*a*) into a series of sines; (*b*) into a series of cosines; (*c*) into a series containing both sines and cosines,— depending on how the function is defined in the interval $-\pi<x<0$.

Which of these developments are uniquely determined?

7. State a generalization of Question 6 for an arbitrary subinterval (a', b'), where $-\pi<a'<b'<\pi$.

2. The General Problem of Development into Series. Power Series. We have met the development of a function into a power series, *Introduction to the Calculus*, Chap. XIV, p. 423:

$$(1) \qquad f(x)=f(a)+f'(a)(x-a)+\frac{f''(a)}{2!}(x-a)^2+\cdots$$

and we have proved that, in the case of some of the most important functions, like e^x, $\sin x$, $\cos x$, $\log(1+x)$, $(1+x)^\mu$, the series converges, at least throughout a certain interval of values for x, and represents the function there. But what reason was there for expecting such a result, and what reason is there for expecting it ever to happen again? Is it not all but preposterous to expect, for example, a series of powers, like

$$x-\frac{x^3}{3!}+\frac{x^5}{5!}-\cdots,$$

whose terms are *not* periodic, and whose partial sum, $s_n(x)$, is not periodic, either, no matter how far n be increased, to represent a periodic function, $\sin x$? For,

$$\sin(x+2\pi)=\sin x.$$

What is there behind it all?

One answer to the question is as follows. Let us give ourselves a succession of polynomials of degrees 0, 1, 2, $\cdots$, n, with undetermined coefficients, and try to determine the latter so as to get the *best approximation possible near the point a.* Write these polynomials as

$$s_n(x)=c_0+c_1(x-a)+c_2(x-a)^2+\cdots+c_{n-1}(x-a)^{n-1}.$$

Thus the graphs of the first three functions:

$$s_1(x)=c_0, \qquad s_2(x)=c_0+c_1(x-a),$$
$$s_3(x)=c_0+c_1(x-a)+c_2(x-a)^2,$$

are recognized as (*i*) a horizontal straight line; (*ii*) an arbitrary straight line (not vertical); (*iii*) a parabola with vertical axis. It is clear that the best use to make of c_0 in Case (*i*) is to make the line go through the point P which corresponds to $x=a$:

$$s_1(x)=f(a);$$

and in Case (*ii*), the best line is the tangent to the curve $y=f(x)$ at P, or

$$s_2(x)=f(a)+f'(a)(x-a).$$

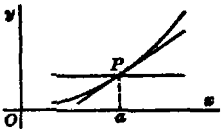


FIG. 102

In Case (*iii*), the best parabola will surely go through P and be tangent there, or

$$s_3(a)=c_0=f(a); \qquad s_3'(a)=c_1=f'(a),$$
$$s_3(x)=f(a)+f'(a)(x-a)+c_2(x-a)^2.$$

Now, of all these parabolas, *that one will most nearly approximate to the given curve, which has the highest contact with it at P*, and this is the one whose curvature is the same. Hence we must have:

$$s_3''(a)=f''(a), \qquad \text{or} \qquad c_2=\tfrac{1}{2}f''(a).$$

Thus

$$s_3(x)=f(a)+f'(a)(x-a)+\frac{f''(a)}{2!}(x-a)^2.$$

The principle is now apparent. We take as our criterion of *best approximation near P* the requirement that the n-th approximation curve have the highest possible order of contact with the given curve at P, and so

$$s_n(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n-1)}(a)}{(n-1)!}(x-a)^{n-1}.$$

In particular, we observe that the coefficients which appear in any given $s_m(x)$ are the same for those respective terms in all later approximations, $s_n(x)$, $n > m$.

If the function $f(x)$ has derivatives of all orders near $x = a$, we can carry this process on indefinitely, and it might seem that the approximation curve, $y = s_n(x)$, must surely, at least throughout a definite interval, approach the given curve, $y = f(x)$, as its limit. But it is to be remembered that a given order of contact becomes, so to speak, geometrically operative only near the point in question; *i.e.* for values of x in a certain interval

$$(2) \qquad a - h < x < a + h.$$

And it is quite conceivable that, as n increases, h should grow smaller and smaller and approach the limit 0. That this does not happen in the case of the most important functions which arise in practice; that, namely, positive *constants* h do exist such that, throughout the whole interval (2), the approximation is uniformly close, — this is one of the great phenomena in mathematics, comparable with the law of gravitation in physics. It bears not only on the character of the functions themselves, but also on the form of the functions of approximation, the particular $s_n(x)$ used; for not every set of functions $s_n(x)$, each of which is itself of the character of a polynomial and has contact of the $(n-1)$-st order at P, has this property.

It can be shown that a power series represents a continuous function throughout its interval of convergence. Hence such functions as those of § 1 which have a discontinuity cannot be represented by a power series throughout the entire interval (a, b).

EXERCISES

1. Plot accurately the graphs of

$$f(x) = \log x \qquad \text{and} \qquad s_n(x)$$

for $a = 1$, $(n = 2, 3, 4, 5)$, to 10 cm. as the unit in the interval $1 \leqq x \leqq 2.1$.

2. Plot the graphs of $\sin x$ and the approximation curves through the one in $x^7/7!$, in the interval $-\pi \leqq x \leqq \pi$, taking $a = 0$.

3. Continuation. Series of Orthogonal Functions. A new approach to the problem of development into series is as follows. Let a function, $f(x)$, be given, which is continuous throughout an interval (a, b), or $a \leqq x \leqq b$; or at least is made up of a finite number of such pieces, as explained in § 1 (cf. Fig. 98). Let

$$(1) \qquad \phi_0(x), \quad \phi_1(x), \quad \phi_2(x), \cdots$$

be a set of standard functions; for example, if the interval (a, b) is $-\pi < x \leqq \pi$, let

$$\phi_{2n}(x) = \cos nx, \quad \phi_{2n-1}(x) = \sin nx, \quad n = 1, 2, \cdots; \quad \phi_0(x) = \tfrac{1}{2}.$$

And let it be proposed to develop $f(x)$ into a series of the form:

$$(2) \qquad f(x) = c_0\phi_0(x) + c_1\phi_1(x) + c_2\phi_2(x) + \cdots.$$

In the case just cited, (2) would be a Fourier's series,

$$(3) \quad f(x) = \tfrac{1}{2}a_0 + a_1\cos x + b_1 \sin x + a_2 \cos 2x + b_2 \sin 2x + \cdots.$$

What sort of functions, $\phi_n(x)$, should we expect to use in the general case, and what kind of requirement should we impose on the approximation curves,

$$(4) \qquad s_n(x) = c_0\phi_0(x) + c_1\phi_1(x) + \cdots + c_{n-1}\phi_{n-1}(x),$$

if the development is to appear plausible?

Orthogonal Functions. The answer to the first question is suggested by the experience of Mathematical Physics, and in particular by the example of Fourier's series. Let $\phi_n(x)$, $n = 0, 1, 2, \cdots$, be continuous in the interval (a, b), and let

$$(5) \qquad \int_a^b \phi_m(x)\,\phi_n(x)\,dx = 0, \qquad m \neq n,$$

no matter how m and n are chosen. Then the functions $\phi_0(x)$, $\phi_1(x) \cdots$, are said to be *orthogonal.* Moreover, we will assume that they are *linearly independent*; cf. Chap. XIV, § 11.

Since no $\phi_n(x)$ vanishes identically, the integral (5) has a positive value when $m = n$. If this value is unity:

$$(6) \qquad \int_a^b [\phi_n(x)]^2\,dx = 1,$$

we say the system of functions $\phi_n(x)$ is *normalized*. An arbitrary system can be reduced to a normalized one by dividing each term $\phi_n(x)$ by the square root (taken with either sign) of the integral (5), formed for $m = n$.

The system of functions

$$(7) \qquad \tfrac{1}{2}, \quad \cos x, \quad \sin x, \quad \cos 2x, \quad \sin 2x, \cdots$$

considered in the interval $(-\pi, \pi)$, forms an orthogonal system; cf. § 1, (4). If, furthermore, each of these functions be divided by $\sqrt{\pi}$, the new system,

$$(8) \qquad \frac{1}{2\sqrt{\pi}}, \quad \frac{\cos x}{\sqrt{\pi}}, \quad \frac{\sin x}{\sqrt{\pi}}, \quad \frac{\cos 2x}{\sqrt{\pi}}, \quad \frac{\sin 2x}{\sqrt{\pi}}, \cdots$$

becomes normalized.

The Formal Development. Let us assume that the given function can be represented by a series of the desired form:

$$(9) \qquad f(x) = c_0\phi_0(x) + c_1\phi_1(x) + c_2\phi_2(x) + \cdots$$

throughout the interval (a, b). If, furthermore, on multiplying (9) through by $\phi_k(x)$, the new series can be integrated term-by-term, we have, by virtue of the orthogonal property (5):

$$(10) \qquad \int_a^b f(x)\,\phi_k(x)\,dx = c_k\int_a^b [\phi_k(x)]^2\,dx.$$

Hence

$$(11) \qquad c_k = \frac{\displaystyle\int_a^b f(x)\,\phi_k(x)\,dx}{\displaystyle\int_a^b [\phi_k(x)]^2\,dx}, \qquad \text{or} \qquad c_k = \int_a^b f(x)\,\phi_k(x)\,dx,$$

in case the ϕ_n's are normalized.

Thus the form of the series is established, in case a development in terms of the ϕ_n's is possible (and, moreover, the integration whereby the c_k's were computed is allowable). It remains to show that the series (9), where the coefficients are given by (11), converges, and that its value is the same as that of the given function, $f(x)$.

4. Approximations according to the Principle of Least Squares
Turning now to the second question, namely, what kind of requirement should be imposed on the approximation curves (4), we find one answer to be that given by Fourier: Determine the coefficients c_k so that $s_n(x)$ will actually coincide with $f(x)$ in n points of the interval. This requirement it is difficult to administer mathematically.

A second form of requirement is suggested by the principle of Least Squares,* and is as follows: Form the integral

$$(12) \qquad u = \int_a^b [f(x) - c_0\phi_0(x) - c_1\phi_1(x) - \cdots - c_{n-1}\phi_{n-1}(x)]^2\,dx.$$

This integral may be considered as the integral of the square of the error, extended over the entire interval. It is a polynomial in the independent variables $c_0, c_1, \cdots, c_{n-1}$, and it is never negative. Moreover, it is evident that it is capable of taking on indefinitely large values for any given n. Hence it must have a minimum value. A necessary condition for a minimum is that

$$\frac{\partial u}{\partial c_k} = -2\int_a^b [f - c_0\phi_0 - \cdots - c_{n-1}\phi_{n-1}]\,\phi_k\,dx = 0.$$

Since the ϕ_i's are orthogonal functions, this equation reduces to the following:

$$(13) \qquad c_k = \frac{\int_a^b f(x)\,\phi_k(x)\,dx}{\int_a^b [\phi_k(x)]^2\,dx}, \qquad \text{or} \qquad c_k = \int_a^b f(x)\,\phi_k(x)\,dx,$$

in case the ϕ_i's are normalized.

Thus we arrive in a most natural manner at the same determination of the c_k's as in the case of the formal development, § 3, but without any assumption concerning the possibility of the actual development. We note, moreover, that here, as in the case of the Fourier development, the c_k's which correspond to a given value of n remain unchanged for all larger values of n.

* Toepler, *Anzeiger der Akad. der Wissenschaften in Wien*, vol. XIII (1876), p. 205. The method goes back to Bessel, *Astronomische Nachrichten* 6 (1828) p. 333.

EXERCISES

1. If $f(x)$ is a function satisfying the conditions of § 1 in the interval (a, b), and if $\phi_0(x), \phi_1(x), \cdots$ are a system of functions, orthogonal and normalized in this interval; if furthermore c_k is determined by (13), show that the integral (12) has the value

$$\int_a^b [f(x)]^2\,dx - c_0^2 - c_1^2 - \cdots - c_{n-1}^2.$$

2. If $c_0, c_1, \cdots$ are determined as in Question 1, show that the series

$$c_0^2 + c_1^2 + c_2^2 + \cdots$$

converges, and that its value does not exceed $\int_a^b [f(x)]^2\,dx.$

(Bessel's Inequality.)

5. Zonal Harmonics.* A further example of a system of orthogonal functions is afforded by *Legendre's Polynomials* (or *Coefficients*), or the *Zonal Harmonics*, $P_n(x)$, considered in the interval $-1 \leqq x \leqq 1$. These are polynomials of degree n (Chap. XIV, § 19):

$$(1)\quad \begin{cases} P_0(x) = 1,\quad P_1(x) = x,\quad P_2(x) = \frac{3}{2}x^2 - \frac{1}{2},\quad P_3(x) = \frac{5}{2}x^3 - \frac{3}{2}x, \cdots \\ P_n(x) = \dfrac{1\cdot 3\cdot 5 \cdot \cdots (2n-1)}{n!}\left\{x^n - \dfrac{n(n-1)}{2(2n-1)}x^{n-2} \right. \\ \qquad\qquad \left. + \dfrac{n(n-1)(n-2)(n-3)}{2\cdot 4\cdot (2n-1)(2n-3)}x^{n-4} + \cdots\right\}. \end{cases}$$

A second formula for $P_n(x)$ is due to *Rodrigues*:

$$(2)\qquad P_n(x) = \frac{1}{2^n n!}\frac{d^n(x^2-1)^n}{dx^n}.$$

These functions have in common with the functions (7), § 3, not only the orthogonal property, but also the following properties: (*i*) they never exceed numerically a certain constant (incidentally, 1); and (*ii*) $P_n(x)$ has n roots in the interval (as over against the $2n$ roots of those functions).

* For a clear and succinct treatment of the properties of these functions, *cf.* Pierpont, *Functions of a Complex Variable*, Chap. XIV, p. 498. Numerous applications are given in Byerly's *Fourier's Series*.

EXERCISES

1. Show that x^n can be expressed linearly in terms of $P_0(x)$, $P_1(x), \cdots, P_n(x)$:

$$x^n = a_0 P_0(x) + a_1 P_1(x) + \cdots + a_{n-1} P_{n-1}(x).$$

Suggestion. Consider first the cases: $n = 0, 1, 2, 3$.

2. With the aid of the method of integration by parts and Rodigues's formula (2), show that

$$\int_{-1}^{1} x^m P_n(x)\,dx = 0, \qquad m < n.$$

Suggestion. Begin with $n = 1, 2$, and 3.

3. From the results of Questions 1 and 2, deduce the orthogonal property of the Polynomials of Legendre:

$$\int_{-1}^{1} P_m(x)\,P_n(x)\,dx = 0 \qquad m \neq n.$$

4. With the aid of the method of integration by parts and Rodigues's formula (2), show that

$$\int_{-1}^{1} [P_n(x)]^2\,dx = \frac{(-1)^n}{2^{2n}(n!)^2}\int_{-1}^{1}(x^2-1)^n \frac{d^{2n}(x^2-1)^n}{dx^{2n}}\,dx$$

$$= \frac{(-1)^n(2n)!}{2^{2n}(n!)^2}\int_{-1}^{1}(x^2-1)^n\,dx.$$

5. With the aid of the method of integration by parts, show that

$$\int_{-1}^{1}(x-1)^n(x+1)^n\,dx = \frac{(-1)^n(n!)^2}{(2n)!}\int_{-1}^{1}(x+1)^{2n}\,dx = \frac{(-1)^n(n!)^2 2^{2n+1}}{(2n)!\,(2n+1)}.$$

6. From the results of Questions 4 and 5, show that

$$\int_{-1}^{1}[P_n(x)]^2\,dx = \frac{2}{2n+1}.$$

7. If $f(x)$ can be developed into a series of zonal harmonics in the interval $(-1, 1)$:

$$f(x) = c_0 P_0(x) + c_1 P_1(x) + c_2 P_2(x) + \cdots,$$

and if the series, after being multiplied by $P_k(x)$, can be integrated term-by-term, show that

$$c_k = \frac{2n+1}{2}\int_{-1}^{1} f(x)\, P_k(x)\, dx.$$

6. Bessel's Functions.* The Bessel's Function, $J_m(x)$, is a certain solution of the linear differential equation of the second order:

$$(1) \qquad \frac{d^2y}{dx^2} + \frac{1}{x}\frac{dy}{dx} + \left(1 - \frac{m^2}{x^2}\right) y = 0.$$

In particular, $J_0(x)$ is that solution of the equation

$$(2) \qquad x\frac{d^2y}{dx^2} + \frac{dy}{dx} + xy = 0, \qquad \text{or} \qquad \frac{d}{dx}\left(x\frac{dy}{dx}\right) + xy = 0$$

which remains finite at the point $x = 0$ and takes on the value 1 there. It is given by the series

$$(3) \qquad J_0(x) = 1 - \frac{x^2}{2^2} + \frac{x^4}{2^2 4^2} - \frac{x^6}{2^2 4^2 6^2} + \cdots.$$

Furthermore,

$$(4) \qquad J_1(x) = \frac{x}{2} - \frac{x^3}{2^2 4} + \frac{x^5}{2^2 4^2 6} - \cdots.$$

Evidently,

$$(5) \qquad J_0'(x) = -J_1(x).$$

It can be shown that these functions have an infinite number of positive roots, each of which is simple. Let those of $J_0(x)$ be denoted by $\lambda_1, \lambda_2, \cdots$.

The functions

$$(6) \qquad \sqrt{x}\, J_0(\lambda_1 x), \qquad \sqrt{x}\, J_0(\lambda_2 x), \qquad \sqrt{x}\, J_0(\lambda_3 x), \cdots$$

when considered in the interval (0, 1), form an orthogonal family, or

$$(7) \qquad \int_0^1 x\, J_0(\lambda_m x)\, J_0(\lambda_n x)\, dx = 0, \qquad m \neq n.$$

For, on making a change of variable, $x = \lambda x'$, $\lambda \neq 0$, we find that a solution of (2) becomes a solution of the new equation

$$x'\frac{d^2y}{dx'^2} + \frac{dy}{dx'} + \lambda^2 x' y = 0.$$

* Cf. Chap. XIV, § 19. The reference to Pierpont in § 5 applies to these functions, too; l. c. Chap. XV, p. 538. — For applications, cf. Byerly, *Fourier's Series*.

Hence $J_0(\lambda x)$ satisfies the equation

$$x\frac{d^2y}{dx^2}+\frac{dy}{dx}+\lambda^2 xy=0. \tag{8}$$

Let $u=J_0(\alpha x)$ and $v=J_0(\beta x)$. Then

$$xu''+u'+\alpha^2 xu=0, \qquad xv''+v'+\beta^2 xv=0.$$

Multiply the first of these equations by $-v$, the second by u, and add:

$$x(uv''-vu'')+uv'-vu'+(\beta^2-\alpha^2)\,xuv=0$$

or

$$\frac{d}{dx}[x(uv'-vu')]+(\beta^2-\alpha^2)\,xuv=0.$$

Hence, integrating and observing that $u'=\alpha J_0'(\alpha x)$, we have:

$$(\beta^2-\alpha^2)\int_0^1 xJ_0(\alpha x)\,J_0(\beta x)=\alpha J_0'(\alpha)\,J_0(\beta)-\beta J_0'(\beta)\,J_0(\alpha). \tag{9}$$

If we set $\alpha=\lambda_m$, $\beta=\lambda_n$, the right-hand side of (9) vanishes, and thus the orthogonal property (7) is proved.

It is furthermore possible to evaluate the integral

$$\int_0^1 x\,[J_0(\lambda_n x)]^2\,dx \tag{10}$$

by means of (9). Differentiate (9) partially with respect to β, and then set $\beta=\alpha$:

$$2\alpha\int_0^1 x[J_0(\alpha x)]^2\,dx=\alpha[J_0'(\alpha)]^2-J_0'(\alpha)\,J_0(\alpha)-\alpha J_0''(\alpha)\,J_0(\alpha).$$

If, now, $\alpha=\lambda_n$, this equation gives:

$$\int_0^1 x[J_0(\alpha x)]^2\,dx=\tfrac{1}{2}[J_0'(\lambda_n)]^2=\tfrac{1}{2}[J_1(\lambda_n)]^2. \tag{11}$$

The Function $J_m(x)$. This function is a solution of equation (1) which, for any positive m (not necessarily an integer) is continuous in the interval $0\leqq x\leqq 1$. It can be written in the form:

$$J_m(x)=x^m\phi(x), \tag{12}$$

where $\phi(x)$ can be expressed as a power series in x^2, and $\phi(0)\neq 0$.

The function $J_m(x)$ has an infinite number of positive roots, each of which is simple.

It can be shown that

(13) $$J'_m(x) = \frac{m}{x} J_m(x) - J_{m+1}(x).$$

EXERCISES

1. Prove that, if the function $f(x)$ can be developed into a series of the form

$$f(x) = c_0 J_0(\lambda_0 x) + c_1 J_0(\lambda_1 x) + c_2 J_0(\lambda_2 x) + \cdots,$$

where $\lambda_0, \lambda_1, \lambda_2, \cdots$ are the positive roots of $J_0(x)$; and if, on multiplying this equation by $xJ_0(\lambda_k x)$, the new series can be integrated term-by-term, then

$$c_k = \frac{2\int_0^1 xf(x) J_0(\lambda_k x)\,dx}{[J_1(\lambda_k)]^2}.$$

2. Show that $J_m(\alpha x)$ satisfies the differential equation:

$$x\frac{d^2y}{dx^2} + \frac{dy}{dx} + \left(\alpha^2 - \frac{m^2}{x^2}\right)xy = 0.$$

3. Prove the relation

$$(\beta^2 - \alpha^2)\int_0^x xJ_m(\alpha x)J_m(\beta x)\,dx$$
$$= x\{\alpha J'_m(\alpha x)J_m(\beta x) - \beta J'_m(\beta x)J_m(\alpha x)\}.$$

4. If $\lambda_0, \lambda_1, \lambda_2, \cdots$ are the positive roots of $J_m(x)$, show that

$$\sqrt{x}\,J_m(\lambda_0 x), \qquad \sqrt{x}\,J_m(\lambda_1 x), \qquad \sqrt{x}\,J_m(\lambda_2 x), \cdots$$

are a system of orthogonal functions.

5. Prove that

$$\int_0^1 x[J_m(\lambda_n x)]^2\,dx = \tfrac{1}{2}[J'_m(\lambda_n)]^2 = \tfrac{1}{2}[J_{m+1}(\lambda_n)]^2.$$

6. If the function $f(x)$ can be developed into a series of the form:

$$f(x) = c_0 J_m(\lambda_0 x) + c_1 J_m(\lambda_1 x) + c_2 J_m(\lambda_2 x) + \cdots$$

in the interval (0, 1), and if, on multiplying this equation by $xJ_m(\lambda_k x)$, the new series can be integrated term-by-term, show that

$$c_k = \frac{2\int_0^1 x f(x)\, J_m(\lambda_k x)\, dx}{[J_{m+1}(\lambda_k)]^2}.$$

CHAPTER XVII

THE CALCULUS OF VARIATIONS AND HAMILTON'S PRINCIPLE

1. Maximum or Minimum of $\int_a^b F\left(x, y, \frac{dy}{dx}\right) dx$. Let $F(x, y, p)$ be a given function of the three independent variables x, y, p. Let two fixed points, A and B, be joined by a curve C:

(1) $$y = f(x),$$

where $f(x)$ and its derivative, $\frac{dy}{dx} = y' = f'(x)$, are continuous in the interval $a \leqq x \leqq b$. Form the integral

(2) $$J = \int_a^b F(x, y, y')\, dx.$$

The value of this integral will depend on the particular curve C. The problem is to find that curve C, *i.e.* that function $f(x)$, for which the integral takes on its least (or its greatest) value.*

It is assumed that $F(x, y, p)$ is continuous, together with its partial derivatives of the first and second order, when (x, y) lies in a given region S, and p has any value whatever. Moreover, the curve C lies in S.

Minimum Surface of Revolution. For example, consider the area, A, of the surface of revolution generated by the rotation of the curve (1) about the axis of x. Since

(3) $$A = 2\pi \int_a^b y\sqrt{1 + \frac{dy^2}{dx^2}}\, dx,$$

* For a comprehensive treatment of the Calculus of Variations, admirably written as regards both accuracy and clearness, the student is referred to two books by Bolza: *Lectures on the Calculus of Variations*, Chicago, 1904; *Vorlesungen über Variationsrechnung*, Leipzig, 1908–09. Cf. also Bliss, *The Calculus of Variations*, Chicago, 1925. The latter book, the first of the Carus Monographs, is intended for the layman and seeks to provide an approach to the subject for those whose mathematical training has not gone beyond the rudiments of the calculus.

it is here a question of finding that curve for which this integral will be least. The function $F(x, y, p)$ is here $y\sqrt{1+p^2}$.

Several Dependent Variables. A more general problem consists in that of joining two fixed points in space by a curve such as will make the integral

$$(4) \qquad J = \int_a^b F(x, y, z, y', z')\,dx$$

a minimum (or a maximum).

2. Euler's Equation. Suppose the problem solved, and let $y = f(x)$ be that function which gives to the integral (2), § 1, its least value — we will restrict ourselves to the case of a *minimum*, for that of a maximum can be reduced to this case by changing the sign of F.

Let $\eta = \phi(x)$ be any function which, together with its first derivative, $\eta' = \phi'(x)$, is continuous in the above interval, $a \leq x \leq b$, and which vanishes at the extremities of the interval:

$$\eta|_{x=a} = 0, \qquad \eta|_{x=b} = 0.$$

Moreover, η shall remain numerically less than a positive constant ϵ, which we choose in advance as small as we please; $|\eta| < \epsilon$.

Form the function

$$Y = y + \eta.$$

Then the value of the integral

$$\int_a^b F(x, Y, Y')\,dx$$

will be at least as great as

$$J_0 = \int_a^b F(x, y, y')\,dx,$$

and the same will be true of each of the integrals

$$(1) \qquad J = \int_a^b F(x, y + \alpha\eta, y' + \alpha\eta')\,dx,$$

where α has any value from -1 to $+1$. It is assumed that η is so chosen that the curve corresponding to $y + \alpha\eta$ lies in S.

This latter integral is a function of α, continuous for the range of values of α considered (cf. Chap. XIX, § 1), and it has its least value when $\alpha = 0$. Since it has a continuous derivative (l. c.) we must have:

$$\left.\frac{dJ}{d\alpha}\right|_{\alpha=0} = 0.$$

Now, by Leibniz's Rule (l. c.),

$$(2) \qquad \frac{dJ}{d\alpha} = \int_a^b \frac{\partial F}{\partial \alpha}\, dx = \int_a^b (\eta F_y + \eta' F_p)\, dx,$$

where the partial derivatives are both formed for the arguments $(x, y + \alpha\eta, y' + \alpha\eta')$. On setting $\alpha = 0$, we find, then:

$$(3) \qquad \int_a^b (\eta F_y + \eta' F_p)\, dx = 0,$$

where F_y and F_p are now formed for the arguments (x, y, y').

This equation must hold for all functions η satisfying the above conditions. But it is not easy to draw inferences from the equation in this form. For that reason we assume that the function y has a continuous second derivative * and proceed to transform the integral of the second term by the method of integration by parts:

$$(4) \qquad \int_a^b \eta' F_p\, dx = \eta F_p \Big|_a^b - \int_a^b \eta \frac{dF_p}{dx}\, dx.$$

The first term on the right disappears, since η vanishes by hypothesis at both extremities of the interval of integration. Hence we have, as the equivalent of (3):

* This requirement involves on its face a restriction of the problem, since it is conceivable that the original problem may have a solution, y, whose second derivative does not exist. Hilbert has, indeed, shown that this cannot be the case. But we are not concerned here with the difficult question of proving mathematically that our problem (after still further restrictions) has one and only one solution, and that this solution satisfies the further demand of possessing a second derivative. On the contrary, we take for granted, in a given case, as more or less plausible from the physical evidence, that the problem will admit a unique solution, continuous together with its derivatives of the first two orders, for we know from experience that problems with a physical pedigree do usually admit the kind of solution expected — and we then turn our efforts to finding this solution. Incidentally, we get some interesting surprises as to what was to be expected from the physical evidence.

$$(5) \qquad \int_a^b \eta\left(F_y - \frac{dF_p}{dx}\right)dx = 0.$$

Now η is arbitrary, subject to the conditions stated above. This fact enables us to infer that, at every point of the interval (a, b), the following equation is satisfied:

$$(A) \qquad F_y - \frac{dF_p}{dx} = 0.$$

For, suppose that, at a point $x = c$ of the interval, the left-hand side were, say, positive. Being a continuous function, it must remain positive throughout a certain neighborhood of c. Let η be so chosen as also to be positive in this neighborhood, but zero everywhere else in the interval. Then the integral could not be zero, and so we have a contradiction. Hence the theorem.

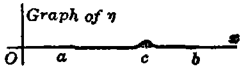


FIG. 103

Equation (A) is known as *Euler's Equation*. It forms a necessary condition for a minimum (or a maximum). On carrying out the differentiation indicated, it takes on the form

$$(A') \qquad F_{pp}\frac{d^2y}{dx^2} + F_{yp}\frac{dy}{dx} + F_{xp} - F_y = 0,$$

a differential equation of the second order for y. The integral of such a differential equation will depend on two arbitrary constants, and these will be determined by the requirement that the curve pass through the fixed points, A and B.

Extremals. A curve corresponding to any solution of Euler's Equation is called an *extremal.*

The Integrand, Independent of x. It may happen that the integrand, $F(x, y, y')$, does not contain x explicitly. In that case a first integral of Euler's Equation can be written down at once. For, since here

$$(6) \qquad \frac{d}{dx}(F - y'F_p) \equiv y'\left(F_y - \frac{d}{dx}F_p\right),$$

we have:

$$(7) \qquad F - y'F_p = \text{const.}$$

Several Dependent Variables. If the integral depends on several functions, as

$$(8) \qquad \int_a^b F(x, y, z, y', z')\,dx,$$

then Euler's Equation must hold for each letter separately:

$$(9) \qquad F_y - \frac{d}{dx} F_{y'} = 0, \qquad F_z - \frac{d}{dx} F_{z'} = 0.$$

On the other hand, the integrand may contain derivatives of highe order than the first. Thus if

$$(10) \qquad J = \int_a^b F(x, y, y', y'')\, dx,$$

and if we consider such functions η as are continuous, together wit their derivatives of the first two orders, and vanish together wit their first derivatives at the end points of the interval, then Euler' Equation becomes:

$$(11) \qquad F_y - \frac{d}{dx} F_{y'} + \frac{d^2}{dx^2} F_{y''} = 0.$$

EXERCISES

1. Let it be required to find the curve which connects two give points in the upper half-plane and makes the integral

$$J = \int_a^b \frac{\sqrt{1 + y'^2}}{y}\, dx$$

a minimum. Show that the extremals are the semicircles whos centres lie on the axis of x, and determine the one which goe through the given points.

2. A ray of light traverses a certain medium, in which its velo ity, $\phi(x, y)$, is variable, but at any given point is the same for a directions. Show that the time required from one fixed point to second is given by the integral:

$$t = \int \frac{ds}{\phi} = \int_a^b \frac{\sqrt{1 + y'^2}}{\phi(x, y)}\, dx,$$

and hence the path must correspond to a solution of the differentia equation:

$$\frac{\phi y''}{1 + y'^2} - \phi_x y' + \phi_y = 0.$$

3. Integrate the differential equation of the preceding questio when

$$\phi(x, y) = cx.$$

4. The differential of arc of a curve on the surface

$$x = f(u, v), \qquad y = \phi(u, v), \qquad z = \psi(u, v)$$

has been shown to be given by the equation (cf. Chap. VI, § 4, Ex. 21)

$$ds^2 = E du^2 + 2F du dv + G dv^2.$$

Show that the geodesics (*i.e.* the shortest lines on the surface) are determined by the differential equation

$$\frac{d}{du}\left(\frac{F + Gv'}{\sqrt{E + 2Fv' + Gv'^2}}\right) = \frac{E_v + 2F_v v' + G_v v'^2}{2\sqrt{E + 2Fv' + Gv'^2}}.$$

3. Minimum Surface of Revolution. Returning to the problem proposed at the end of § 1, we see that Euler's Equation here becomes

(1) $$\sqrt{1+p^2} - \frac{d}{dx}\frac{yp}{\sqrt{1+p^2}} = 0.$$

Observing that $\frac{d}{dx} = p\frac{d}{dy}$ and performing the differentiations indicated, we have, on reducing the result,

$$yp\frac{dp}{dy} = 1 + p^2.$$

Hence

$$\frac{dy}{y} = \frac{p\,dp}{1+p^2},$$

$$\log y = \tfrac{1}{2}\log(1+p^2) + C,$$

(2) $$\sqrt{1+p^2} = \frac{y}{b},$$

where we have replaced the constant of integration, C, by the equally arbitrary constant, b, setting $C = \log b$.

From (2) it follows that

$$p = \frac{dy}{dx} = \pm\sqrt{\frac{y^2}{b^2} - 1}.$$

Hence

$$dx = \pm\frac{b\,dy}{\sqrt{y^2 - b^2}},$$

$$x - a = \pm b \log\frac{y + \sqrt{y^2 - b^2}}{b}.$$

On solving this equation for y, we find:

$$y = \frac{b}{2}\left(e^{\frac{x-a}{b}} + e^{-\frac{x-a}{b}}\right),$$

the equation of a catenary, referred to the axis of x as directrix. The scale to which the curve is drawn is arbitrary, as shown by the constant b; and the location of the axis of y is also arbitrary, as shown by the constant a. But the directrix is fixed.

To complete the solution of our problem, we must choose a and b so that the catenary will go through the two fixed points. Can this always be done? There is evidence from mechanics on this point; cf. Chap. XIV, § 7. Let a heavy flexible string be hung over two smooth pegs at the fixed points, and let its ends reach down to the directrix, D, of the particular catenary in which it hangs. Then it can be released at the fixed points and will not slip. If D, perchance, coincides with the axis of x, our solution is complete; the part of the string between the pegs gives us the desired catenary.

If D, on the other hand, lies above the axis of x, then a steady lengthening of the string will bring it down, and when it reaches the axis of x, our problem is solved.

If, however, D lies below the axis of x, it may be possible to raise D by shortening the string; but there is a limit, above which D cannot go, cf. Chap. XIV, § 7, Ex. 1, and if this limit is below the axis of x, our problem cannot be solved.

What does this mean for the minimum problem with which we started? Will there not always be a surface of revolution of least area? No! If the fixed points are very near the axis of revolution, in comparison with the distance between them, a surface of revolution generated by a curve like the one indicated will have a relatively small area, and these various areas will have a certain *lower limit*, which can be shown to be the sum of the areas of the two discs, whose radii are y_0 and y_1. But evidently no surface of revolution can be found which will have quite so small an area.

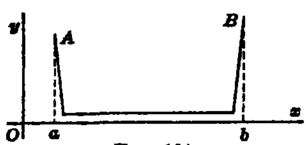


FIG. 104

Thus it is seen that a problem in the Calculus of Variations, which it is easy to formulate, may not have a solution.

4. The Brachystochrone. Let it be required to find the curve of quickest descent, *i.e.* the form for a wire on which a smooth bead is to slide, in order that the bead, leaving O from rest, may arrive at A in the shortest possible time.

Choosing the axes as indicated, we have, for an arbitrary point of the path (cf. *Introduction to the Calculus*, p. 375):

$$v^2 = 2gy; \qquad v = \frac{ds}{dt},$$

$$t = \frac{1}{\sqrt{2g}} \int_0^l \frac{ds}{\sqrt{y}},$$

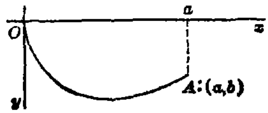


FIG. 105

and so it is a question of making the integral

$$(1) \qquad \int_0^a \frac{\sqrt{1+y'^2}}{\sqrt{y}}\,dx$$

a minimum.

Euler's Equation here becomes

$$(2) \qquad \frac{\sqrt{1+p^2}}{2y^{3/2}} + \frac{d}{dx}\frac{p}{\sqrt{y}\sqrt{1+p^2}} = 0.$$

This reduces to the equation:

$$1 + \frac{2yp}{1+p^2}\frac{dp}{dy} = 0.$$

Integrating, we have:

$$1 + p^2 = \frac{2c}{y}.$$

Hence
$$\frac{dy}{dx} = \sqrt{\frac{2c-y}{y}},$$

$$(3) \qquad x = \int \frac{y\,dy}{\sqrt{2cy-y^2}} = -\sqrt{2cy-y^2} + c\cos^{-1}\frac{c-y}{c} + C.$$

Since the curve starts at the origin, $C = 0$.

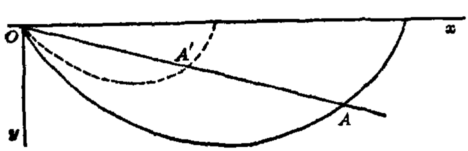


FIG. 106

This is the equation of a cycloid,

$$(4) \qquad x = c\theta - c\sin\theta,$$
$$y = c - c\cos\theta.$$

To determine c, draw the particular cycloid, for which $c = 1$, and let the line OA (produced, if necessary) cut it in A'. Then, since all cycloids are similar, $c = \overline{OA}/\overline{OA'}$.

5. Definition of the Variations. The arbitrary function η of § 2 is an illustration of what is known as a *variation*, and since it is added to y, it is denoted by δy and read: "variation of y,"

(1) $$\eta = \delta y.$$

Its derivatives are denoted as the *variations of the derivatives of* y,

(2) $$\eta' = \delta y', \qquad \eta'' = \delta y'', \qquad \text{etc.}$$

Thus by definition

(3) $$\frac{d\,\delta y}{dx} = \delta \frac{dy}{dx}.$$

Variation of a Function, $F(x, y, y', y'', \cdots z, z', \cdots)$. Let F be a function of $x, y, y', y'', \cdots, z, z', \cdots$, continuous together with its first partial derivatives, and let $y, z, \cdots$ be functions of the independent variable x, continuous together with such of their derivatives as enter in F. Let $y, z, \cdots$ receive variations $\delta y, \delta z, \cdots$. The variation of F is defined as follows:

(4) $$\delta F = F_y \delta y + F_{y'} \delta y' + F_{y''} \delta y'' + \cdots + F_z \delta z + F_{z'} \delta z' + \cdots.$$

Here, $\delta y, \delta z, \cdots$ are chosen arbitrarily, subject to such conditions as those we met in § 2.

The definition applies equally well in case there are several independent variables.

The independent variable or variables are not varied. Thus

(5) $$\delta F\left(x, y, u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}\right) = F_u \delta u + F_p \delta p + F_q \delta q,$$

$$p = \frac{\partial u}{\partial x}, \qquad q = \frac{\partial u}{\partial y},$$

where $\zeta = \delta u$ is an arbitrary function, continuous together with its partial derivatives

(6) $$\zeta_x = \frac{\partial \zeta}{\partial x} = \delta p, \qquad \zeta_y = \frac{\partial \zeta}{\partial y} = \delta q,$$

throughout the region S in which u is considered, and required to be numerically small.

If $$J = \int_a^b F(x, y, y')\, dx,$$

then δJ is defined as follows:

$$(7) \qquad \delta J = \int_a^b (F_y \,\delta y + F_{y'} \,\delta y')\, dx.$$

Thus

$$(8) \qquad \delta J = \int_a^b \delta F\, dx, \qquad \text{or} \qquad \delta \int_a^b F\, dx = \int_a^b \delta F\, dx.$$

It may happen, however, that δy is restricted so that a second integral will have a constant value:

$$(9) \qquad K = \int_a^b \Phi(x, y + \delta y, y' + \delta y')\, dx.$$

This can be accomplished by making δy depend on a parameter, as is done below in § 8.

For other forms of definition of variations, cf. Bolza, *Variationsrechnung*, p. 45, § 8.

6. Euler's Equation for Multiple Integrals. Consider the integral *

$$(1) \qquad J = \iint_S F\left(x, y, u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}\right) dS.$$

Let u be the function which makes J a minimum. Give to u a variation, $\delta u = \zeta(x, y)$, where ζ vanishes on the boundary, C, of S:

$$(2) \qquad \zeta|_C = 0.$$

Then the function of α,

$$(3) \qquad J(\alpha) = \iint_S F(x, y, u + \alpha\zeta, p + \alpha\zeta_x, q + \alpha\zeta_y)\, dS,$$

has a minimum when $\alpha = 0$, and $J'(0) = 0$. Now, by Leibniz's Rule,

$$(4) \qquad \frac{dJ}{d\alpha} = \iint_S (\zeta F_u + \zeta_x F_p + \zeta_y F_q)\, dS,$$

* As to continuity it is assumed: (i) that $F(x, y, u, p, q)$ is continuous, together with its partial derivatives of the first two orders, when (x, y, u) lies in a given region, V, of space, and p, q have any values whatever. Moreover, the surface $u = f(x, y)$, for which the integral is formed, shall lie in V, and $f(x, y)$ shall be continuous, together with its derivatives of the first two orders, in S. The same shall hold true for the varied surfaces, $u + \alpha\zeta$.

where F_u, F_p, F_q are formed for the arguments $(x, y, u + \alpha\zeta, p + \alpha\zeta_x, q + \alpha\zeta_y)$. On setting $\alpha = 0$, we have, therefore,

(5) $$\iint_S (\zeta F_u + \zeta_x F_p + \zeta_y F_q)\, dS = 0,$$

where these partial derivatives are formed for the arguments (x, y, u, p, q).

This equation corresponds to equation (3) of § 2, and, like that one, can be transformed by integration by parts. Since *

$$\frac{\partial}{\partial x}(\zeta F_p) = \zeta_x F_p + \zeta \frac{\partial F_p}{\partial x},$$

we have:

$$\iint_S \zeta_x F_p\, dS = \int_C \zeta F_p\, dy - \iint_S \zeta \frac{\partial F_p}{\partial x}\, dS.$$

But ζ vanishes by hypothesis on the boundary, and so the line integral drops out.

On transforming the last integral in (5) by a similar method we find, as the equivalent of (5):

(6) $$\iint_S \zeta\left(F_u - \frac{\partial F_p}{\partial x} - \frac{\partial F_q}{\partial y}\right) dS = 0.$$

Now ζ is arbitrary, within the limits imposed. We can, therefore, infer that equation (6) is true in all cases only when the second factor in the integrand vanishes at every point of S, or

(7) $$F_u - \frac{\partial F_p}{\partial x} - \frac{\partial F_q}{\partial y} = 0.$$

This is *Euler's Equation.* In it, F_p and F_q are formed for the arguments $x, y, u, \partial u/\partial x, \partial u/\partial y$, where u denotes the function that makes J a minimum. Thus F_p and F_q become functions of the two independent variables, x and y, and the differentiations indicated by the ∂'s are performed under this hypothesis.

Generalized Coordinates. It should be observed that, although equation (5) holds for any system of coordinates whatever, equation (7), depending as it does on integration by parts, is restricted to the

* This is the transformation known as Green's Theorem; cf. Chap. XI, § 9.

case of *Cartesian coordinates*. If curvilinear coordinates (λ, μ) be introduced, then (Chap. XII, § 4)

$$J = \pm \iint_{\Sigma} \Phi(\lambda, \mu, u, \rho, \sigma)\, D\, d\Sigma, \tag{8}$$

where
$$F\left(x, y, u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}\right) = \Phi\left(\lambda, \mu, u, \frac{\partial u}{\partial \lambda}, \frac{\partial u}{\partial \mu}\right),$$

$$D = \frac{\partial(x, y)}{\partial(\lambda, \mu)}, \qquad \rho = \frac{\partial u}{\partial \lambda}, \qquad \sigma = \frac{\partial u}{\partial \mu},$$

and the integral is extended over that region, Σ, of the Cartesian (λ, μ)-plane, into which S is transformed by the equations which correspond to the system of curvilinear coordinates:

$$\lambda = f(x, y), \qquad \mu = \phi(x, y). \tag{9}$$

Thus Euler's Equation takes on the form:

$$D\Phi_u - \frac{\partial}{\partial \lambda}(D\Phi_\rho) - \frac{\partial}{\partial \mu}(D\Phi_\sigma) = 0. \tag{10}$$

EXERCISES

1. If u is a function which renders the integral

$$\iint_S \left\{ \frac{\partial u^2}{\partial x^2} + \frac{\partial u^2}{\partial y^2} \right\} dS$$

a minimum, show that

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0.$$

2. If u is a function which renders the integral

$$\iint_S \left\{ \left(\frac{\partial u}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial u}{\partial \theta}\right)^2 \right\} dS$$

a minimum, r and θ being polar coordinates, show that

$$r^2 \frac{\partial^2 u}{\partial r^2} + r\frac{\partial u}{\partial r} + \frac{\partial^2 u}{\partial \theta^2} = 0.$$

3. Show that Euler's Equation for the volume integral

$$J = \iiint_V F(x, y, z, u, p, q, r)\, dV,$$

where
$$p=\frac{\partial u}{\partial x}, \qquad q=\frac{\partial u}{\partial y}, \qquad r=\frac{\partial u}{\partial z},$$
is:
$$F_u-\frac{\partial F_p}{\partial x}-\frac{\partial F_q}{\partial y}-\frac{\partial F_r}{\partial z}=0.$$

Explain carefully the meaning of each partial derivative.

4. If u is a function which renders the integral
$$\iiint_V \left\{\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial u}{\partial y}\right)^2+\left(\frac{\partial u}{\partial z}\right)^2\right\} dV$$
a minimum, show that
$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}+\frac{\partial^2 u}{\partial z^2}=0.$$

5. If u is a function which renders the integral
$$\iiint_V \left\{\left(\frac{\partial u}{\partial r}\right)^2+\frac{1}{r^2}\left(\frac{\partial u}{\partial \phi}\right)^2+\frac{1}{r^2\sin^2\phi}\left(\frac{\partial u}{\partial \theta}\right)^2\right\} dV$$
a minimum, where (r, ϕ, θ) are spherical polar coordinates, show that
$$\frac{\partial}{\partial r}\left(r^2\sin\phi\,\frac{\partial u}{\partial r}\right)+\frac{\partial}{\partial \phi}\left(\sin\phi\,\frac{\partial u}{\partial \phi}\right)+\frac{1}{\sin\phi}\,\frac{\partial^2 u}{\partial \theta^2}=0.$$

6. *Minimum Surface.* A surface, $z=f(x, y)$, is spanned into a simple closed twisted curve. Show that, if its area
$$\iint_S \sqrt{1+p^2+q^2}\,dS$$
is to be a minimum, then
$$(1+q^2)\frac{\partial^2 z}{\partial x^2}-2pq\,\frac{\partial^2 z}{\partial x\,\partial y}+(1+p^2)\,\frac{\partial^2 z}{\partial y^2}=0.$$

7. If the integrand of the double integral contains the second derivatives,
$$r=\frac{\partial^2 u}{\partial x}, \qquad s=\frac{\partial^2 u}{\partial x\,\partial y}, \qquad t=\frac{\partial^2 u}{\partial y^2},$$
so that we have:
$$J=\iint_S F(x, y, u, p, q, r, s, t)\,dS,$$

show that Euler's Equation becomes

$$F_u - \frac{\partial F_p}{\partial x} - \frac{\partial F_q}{\partial y} + \frac{\partial^2 F_r}{\partial x^2} + \frac{\partial^2 F_s}{\partial x \partial y} + \frac{\partial^2 F_t}{\partial y^2} = 0.$$

8. Extend Ex. 7 to volume integrals.

7. Laplace's Equation in Curvilinear Coordinates. Dirichlet's Principle. A classical problem of mathematical physics is the following, known as the (first) *boundary value problem* for Laplace's equation: To find a function, u, which throughout a given region V of space satisfies Laplace's equation,

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0,$$

and which, moreover, takes on given boundary values along the boundary of V.

The mathematicians of the middle of the nineteenth century sought to solve the problem of showing that such a function always exists by formulating a problem in the calculus of variations, which is identical with the original problem, and the answer to which seemed to them self-evident. They considered, namely, the integral

$$(1) \qquad \iiint\limits_V \left\{ \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial u}{\partial z}\right)^2 \right\} dV,$$

extended throughout the given region, the function u being required to be continuous, together with its derivatives of the first two orders, and to take on the given boundary values. Euler's Equation for this problem is (§ 6, Ex. 4) precisely Laplace's equation. Now, the integrand of the integral (1) is never negative, and so the value of the integral cannot be negative. It is evident, therefore, that the value of the integral, corresponding to various choices of u, has a *lower limit* which is not negative; *i.e.* that there exists a constant $A \geqq 0$ such that, if ϵ be as small a positive constant as you please, there will be some function u for which the value of the integral will be less than $A + \epsilon$. But how do we know that there is a function u for which the value of the integral *reaches* this lower limit, A? We have seen, in the case of the minimum surface of revolution, that a perfectly well appearing problem of the calculus of variations may have no solution. A lower limit may exist when a minimum does not. What is the lower limit of the positive numbers? What is the smallest positive number?

Weierstrass pointed out the fallacy in the assumption that there must be a function u for which the integral reaches its lower limit, and thus this method of proving the existence theorem for Laplace's equation — the method known as *Dirichlet's Principle* — fell to the ground.

If, on the other hand, we have a solution, u, of Laplace's equation, it can be shown that this function makes the integral (1) a minimum, and hence the problem of the calculus of variations does have a solution. For, let $U = u + h$ be any other function of the class admitted to competition, and form the integral (1) for U. Thus we find:

$$\iiint_V \left\{\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial u}{\partial z}\right)^2\right\} dV$$
$$+ 2\iiint_V \left\{\frac{\partial h}{\partial x}\frac{\partial u}{\partial x} + \frac{\partial h}{\partial y}\frac{\partial u}{\partial y} + \frac{\partial h}{\partial z}\frac{\partial u}{\partial z}\right\} dV$$
$$+ \iiint_V \left\{\left(\frac{\partial h}{\partial x}\right)^2 + \left(\frac{\partial h}{\partial y}\right)^2 + \left(\frac{\partial h}{\partial z}\right)^2\right\} dV.$$

The value of the second integral is seen by Green's Theorem to be 0; cf. Chap. XI, § 9, Ex. 2, in which the u and v are to be replaced by the present h and u respectively, and observe that h vanishes on the boundary.

The value of the last integral is positive, unless $h = \text{const.}$ But since $h = 0$ on the boundary, the constant would be zero, and thus U would $= u$.

Laplace's Equation. This result is important in practice, since it enables us to simplify the computation of Laplace's equation in curvilinear coordinates. Let

$$(2) \qquad \lambda = f(x, y, z), \qquad \mu = \phi(x, y, z), \qquad \nu = \psi(x, y, z)$$

be the equations of three families of surfaces (λ, μ, ν being parameters), and let each surface of any one family cut each surface of any other family orthogonally. The relations existing between the partial derivatives of these functions and those of the inverse functions are developed in Chap. VI, § 6, and it is there shown in particular that

$$(3) \quad \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial u}{\partial z}\right)^2 = \frac{1}{H_1}\left(\frac{\partial u}{\partial \lambda}\right)^2 + \frac{1}{H_2}\left(\frac{\partial u}{\partial \mu}\right)^2 + \frac{1}{H_3}\left(\frac{\partial u}{\partial \nu}\right)^2,$$

$$(4) \qquad J = \frac{\partial(x, y, z)}{\partial(\lambda, \mu, \nu)} = \pm\sqrt{H_1 H_2 H_3}.$$

Thus the integral (1), when transformed to the curvilinear coordinates λ, μ, ν, becomes (save as to sign)

$$(5) \qquad \iiint_{\tau} \left\{ \frac{1}{H_1}\left(\frac{\partial u}{\partial \lambda}\right)^2 + \frac{1}{H_2}\left(\frac{\partial u}{\partial \mu}\right)^2 + \frac{1}{H_3}\left(\frac{\partial u}{\partial \nu}\right)^2 \right\} \sqrt{H_1 H_2 H_3}\, d\tau,$$

where τ denotes the region of the (λ, μ, ν)-space which corresponds to the given region V of the (x, y, z)-space. Since a function, u, which satisfies Laplace's equation, makes the integral (1) a minimum, such a function must also make the integral (5) a minimum.

Euler's Equation now takes on the form:

$$(6) \qquad \frac{\partial}{\partial \lambda}\left(\sqrt{\frac{H_2 H_3}{H_1}}\frac{\partial u}{\partial \lambda}\right) + \frac{\partial}{\partial \mu}\left(\sqrt{\frac{H_3 H_1}{H_2}}\frac{\partial u}{\partial \mu}\right) + \frac{\partial}{\partial \nu}\left(\sqrt{\frac{H_1 H_2}{H_3}}\frac{\partial u}{\partial \nu}\right) = 0,$$

and this is Laplace's equation in orthogonal curvilinear coordinates.*

Example. Let the curvilinear coordinates be spherical polar coordinates:

$$x = r \sin \phi \cos \theta, \qquad y = r \sin \phi \sin \theta, \qquad z = r \cos \phi.$$

Then $\qquad H_1 = 1, \qquad H_2 = r^2, \qquad H_3 = r^2 \sin^2 \phi;$

$$(8) \qquad \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial u}{\partial z}\right)^2 = \left(\frac{\partial u}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial u}{\partial \phi}\right)^2 + \frac{1}{r^2 \sin^2 \phi}\left(\frac{\partial u}{\partial \theta}\right)^2;$$

$$J = r^2 \sin \phi.$$

Laplace's equation becomes:

$$(9) \qquad \frac{\partial}{\partial r}\left(r^2 \sin \phi \frac{\partial u}{\partial r}\right) + \frac{\partial}{\partial \phi}\left(\sin \phi \frac{\partial u}{\partial \phi}\right) + \frac{1}{\sin \phi}\frac{\partial^2 u}{\partial \theta^2} = 0.$$

* For a treatment purely by partial differentiation, in which, moreover, the identity

$$(7) \quad \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = \frac{1}{\sqrt{H_1 H_2 H_3}}\left[\frac{\partial}{\partial \lambda}\left(\sqrt{\frac{H_2 H_3}{H_1}}\frac{\partial u}{\partial \lambda}\right) + \frac{\partial}{\partial \mu}\left(\sqrt{\frac{H_3 H_1}{H_2}}\frac{\partial u}{\partial \mu}\right) + \frac{\partial}{\partial \nu}\left(\sqrt{\frac{H_1 H_2}{H_3}}\frac{\partial u}{\partial \nu}\right)\right]$$

is established, cf. Goursat-Hedrick, *Mathematical Analysis*, vol. I, Chap. II, § 43. The more general case of the Exercise below is treated on the basis of the vanishing of the variation, and not of the minimum property, by Courant and Hilbert, *Methoden der mathematischen Physik*, vol. I, p. 194. Why should the vanishing of the variation be independent of the choice of the coordinates?

The proof given in the text is in one respect incomplete, since the *uniqueness* of the equation (6) has not been established.

EXERCISE

Let the requirement be removed from the transformation (2) that it be orthogonal. Then

$$(10)\quad \left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial u}{\partial y}\right)^2+\left(\frac{\partial u}{\partial z}\right)^2=A_1\left(\frac{\partial u}{\partial \lambda}\right)^2+A_2\left(\frac{\partial u}{\partial \mu}\right)^2+A_3\left(\frac{\partial u}{\partial \nu}\right)^2$$
$$+2B_1\frac{\partial u}{\partial \mu}\frac{\partial u}{\partial \nu}+2B_2\frac{\partial u}{\partial \nu}\frac{\partial u}{\partial \lambda}+2B_3\frac{\partial u}{\partial \lambda}\frac{\partial u}{\partial \mu},$$

where $\qquad A_1=\Delta_1\lambda, \qquad A_2=\Delta_1\mu, \qquad A_3=\Delta_1\nu,$

$$B_1=\Delta(\mu,\nu), \qquad B_2=\Delta(\nu,\lambda), \qquad B_3=\Delta(\lambda,\mu),$$

the notation $\Delta(\lambda,\mu)$ standing for the first polar of $\Delta_1\lambda$ with respect to $\left(\frac{\partial \mu}{\partial x}, \frac{\partial \mu}{\partial y}, \frac{\partial \mu}{\partial z}\right)$, or

$$\Delta(\lambda,\mu)=\frac{\partial \lambda}{\partial x}\frac{\partial \mu}{\partial x}+\frac{\partial \lambda}{\partial y}\frac{\partial \mu}{\partial y}+\frac{\partial \lambda}{\partial z}\frac{\partial \mu}{\partial z}.$$

Moreover,

$$j^2=\begin{vmatrix} A_1 & B_3 & B_2 \\ B_3 & A_2 & B_1 \\ B_2 & B_1 & A_3 \end{vmatrix}, \qquad j=\frac{\partial(\lambda,\mu,\nu)}{\partial(x,y,z)}=\frac{1}{J}, \qquad J=\frac{\partial(x,y,z)}{\partial(\lambda,\mu,\nu)}.$$

Show that Laplace's equation is the following:

$$\frac{\partial}{\partial \lambda}\left[\left(A_1\frac{\partial u}{\partial \lambda}+B_3\frac{\partial u}{\partial \mu}+B_2\frac{\partial u}{\partial \nu}\right)J\right]$$
$$+\frac{\partial}{\partial \mu}\left[\left(B_3\frac{\partial u}{\partial \lambda}+A_2\frac{\partial u}{\partial \mu}+B_1\frac{\partial u}{\partial \nu}\right)J\right]$$
$$+\frac{\partial}{\partial \nu}\left[\left(B_2\frac{\partial u}{\partial \lambda}+B_1\frac{\partial u}{\partial \mu}+A_3\frac{\partial u}{\partial \nu}\right)J\right]=0.$$

8. Isoperimetric Problems. Consider the problem of joining two points by a curve of given length, so chosen that the area enclosed between the curve and its chord will be a maximum. The area is given by the integral

$$(1)\qquad J=\int_a^b y\,dx;$$

FIG. 107

the length of the curve is

$$(2)\qquad l=\int_a^b\sqrt{1+y'^2}\,dx.$$

And now not all functions y which correspond to curves joining the two fixed points may be admitted to competition, but only such as make the second integral equal to l.

The foregoing example is typical for the general case. Let $F(x, y, p)$ and $\Phi(x, y, p)$ be two functions which, together with their derivatives of the first two orders, are continuous, where (x, y) is any point of a given region S, and p has any value whatever. Form the integrals:

$$(A)\qquad J=\int_a^b F(x, y, y')\,dx, \qquad K=\int_a^b \Phi(x, y, y')\,dx,$$

the curve $y=f(x)$ lying in S. To find that function, y, of x which will make the first integral a minimum (or a maximum), while giving to the second integral a preassigned fixed value, K.

We proceed to deduce a necessary condition in the form of Euler's Equation for the present case. Let η and ζ be two functions which satisfy the conditions imposed on η in § 2. Let us try to determine β as a function of α such that

$$(3)\qquad \Omega(\alpha, \beta)=\int_a^b \Phi(x, y+\alpha\eta+\beta\zeta,\ y'+\alpha\eta'+\beta\zeta')\,dx-K$$

will always $=0$. Here, $\Omega(\alpha, \beta)$ is a continuous function of the two independent variables, α and β; moreover, it has a continuous partial derivative with respect to β (Chap. XIX, § 1),

$$(4)\qquad \frac{\partial\Omega}{\partial\beta}=\int_a^b (\zeta\Phi_y+\zeta'\Phi_p)\,dx,$$

where Φ_y and Φ_p are formed for the arguments

$$(x, y+\alpha\eta+\beta\zeta,\ y'+\alpha\eta'+\beta\zeta').$$

If this derivative, $\Omega_\beta(\alpha, \beta)$, does not vanish at the origin, then the theorem on implicit functions, Chap. V, § 12, tells us that there exists a function, $\phi(\alpha)$, continuous in the neighborhood of the origin and vanishing there, such that the equation

$$(5)\qquad \Omega(\alpha, \beta)=0$$

will be satisfied if $\beta=\phi(\alpha)$.

Now, $\Omega_\beta(0, 0)$ is given by (4), where Φ_y and Φ_p are to be formed for the arguments (x, y, y'). The integral of the second term can be transformed as in § 2, and hence we have:

(6) $$\Omega_\beta(0,\,0)=\int_a^b\zeta\left(\Phi_y-\frac{d\Phi_p}{dx}\right)dx.$$

It is obviously possible to choose ζ so that this last integral will not vanish unless the second factor in the integrand, $\Phi_y - d\Phi_p/dx$, vanishes identically. As a matter of fact, this case does not arise in practice, and so we choose ζ, once for all, so that $\Omega_\beta(0,\,0)\neq 0$.

We can now choose η arbitrarily and then, for any α numerically not too large, determine β so that $\Omega(\alpha,\,\beta)=0$; *i.e.* the second integral in (A) has the required value.

The first integral in (A), formed for the function $Y=y+\alpha\eta+\beta\zeta$, is a function of α which has a minimum when $\alpha=0$. Hence $(dJ/d\alpha)_{\alpha=0}=0$. Now,

(7) $$\frac{dJ}{d\alpha}=\int_a^b\{(\eta+\beta'\zeta)F_y+(\eta'+\beta'\zeta')F_p\}\,dx$$

$$=\int_a^b(\eta F_y+\eta' F_p)\,dx+\beta'\int_a^b(\zeta F_y+\zeta' F_p)\,dx,$$

where F_y and F_p are formed for the arguments $(x,\,y+\alpha\eta+\beta\zeta,\,y'+\alpha\eta'+\beta\zeta')$. Let $\alpha=0$; then $\beta=0$, and since

(8) $$\beta'=-\frac{\Omega_\alpha(\alpha,\,\beta)}{\Omega_\beta(\alpha,\,\beta)},$$

we have:

(9) $$\beta'|_{\alpha=0}=-\frac{\displaystyle\int_a^b(\eta\Phi_y+\eta'\Phi_p)\,dx}{\displaystyle\int_a^b(\zeta\Phi_y+\zeta'\Phi_p)\,dx}.$$

Substitute this value of β' in (7), written for $\alpha=0$. If we set

(10) $$\lambda=-\frac{\displaystyle\int_a^b(\zeta F_y+\zeta' F_p)\,dx}{\displaystyle\int_a^b(\zeta\Phi_y+\zeta'\Phi_p)\,dx},$$

the new equation becomes:

$$\int_a^b (\eta F_y + \eta' F_p)\,dx + \lambda \int_a^b (\eta \Phi_y + \eta' \Phi_p)\,dx = 0,$$

or

(11)
$$\int_a^b \left\{ \eta \overline{F_y + \lambda\Phi_y} + \eta' \overline{F_p + \lambda\Phi_p} \right\} dx = 0.$$

Now this is precisely equation (3) of § 2, written for the function $F + \lambda\Phi$ instead of F. It can be transformed as before by integration by parts, and thus we find:

(12)
$$(F_y + \lambda\Phi_y) - \frac{d}{dx}(F_p + \lambda\Phi_p) = 0.$$

This is Euler's Equation in the parametric case. It expresses a necessary condition for a minimum (or maximum) in terms of an unknown constant, λ. The extremals depend on λ, and on the two further constants of integration.

Example. Recurring to the example with which the paragraph opened, we have:

$$F + \lambda\Phi = y + \lambda\sqrt{1 + p^2}.$$

Thus Euler's Equation becomes:

$$1 - \frac{d}{dx}\frac{\lambda p}{\sqrt{1 + p^2}} = 0.$$

Hence
$$\frac{\lambda p}{\sqrt{1 + p^2}} = x - c.$$

Thus
$$p^2 = \frac{(x - c)^2}{\lambda^2 - (x - c)^2};$$

$$y = \pm \int \frac{(x - c)\,dx}{\sqrt{\lambda^2 - (x - c)^2}} = \mp \sqrt{\lambda^2 - (x - c)^2} + d,$$

or

$$(x - c)^2 + (y - d)^2 = \lambda^2.$$

The extremal is, therefore, the arc of a circle joining the two points and having the prescribed length. The two constants of integration, c and d, and the constant λ of Euler's Equation are just sufficient to permit the fulfilment of the above conditions.

EXERCISES

1. Two fixed points are joined by a uniform heavy wire of given length, which may be bent into any shape. Assuming that there is a form of the wire for which the centre of gravity is lowest, show that this form is the catenary in the vertical plane through the given points.* The extremals are the curves

$$y+\lambda=\frac{c}{2}\left(e^{\frac{x-a}{c}}+e^{-\frac{x-a}{c}}\right).$$

2. Extend the method of the text to double integrals, and show that, if

$$J=\int_S\!\!\int F(x,\,y,\,u,\,p,\,q)\,dS, \qquad K=\int_S\!\!\int \Phi(x,\,y,\,u,\,p,\,q)\,dS,$$

where K is to remain constant, then a necessary condition that J be a maximum or a minimum is given in the form of Euler's Equation:

$$\overline{F_u+\lambda\Phi_u}-\frac{\partial}{\partial x}\overline{F_p+\lambda\Phi_p}-\frac{\partial}{\partial y}\overline{F_q+\lambda\Phi_q}=0.$$

3. Let S be a region of the (x,y)-plane, bounded by a simple closed curve, C; and let a cylinder Z be erected on C with its elements parallel to the axis of z. Let a simple closed (twisted) curve $\mathfrak{C}$ be drawn on Z above the (x,y)-plane. To find the surface of least area which, together with S and the part of Z between C and $\mathfrak{C}$, will enclose a given volume.

Ans. If $z=f(x,\,y)$ is the equation of the surface, Euler's Equation becomes:

$$\frac{(1+q^2)\,r-2pq\,s+(1+p^2)\,t}{(1+p^2+q^2)^{\frac{3}{2}}}=\lambda,$$

where

$$r=\frac{\partial^2 z}{\partial x^2},\qquad s=\frac{\partial^2 z}{\partial x\partial y},\qquad t=\frac{\partial^2 z}{\partial y^2}.$$

The geometric meaning of the result is that the surface is one of constant mean curvature. The surface can be realized physically by means of a soap bubble film spanned into $\mathfrak{C}$, sufficient air being

* The result is altogether reasonable. For, if the wire be thought of as a wet string, frozen in a particular shape, and then allowed to thaw, the string will then change its shape (unless it was a catenary to begin with) and this means precisely to our mechanical intuition that the centre of gravity has fallen. So no shape but the catenary can yield a minimum.

forced into the closed region above mentioned (or withdrawn from it) to yield the given volume. Cf. Bolza, *Variationsrechnung*, p. 662.

9. Variable End Points. Let it be required to find the minimum surface of revolution when the generating curve is to have one end point, A, fixed and the other, B, on a fixed circle. It is clear that the curve must be a catenary; but it is not clear at what point, B, it meets the given circle.

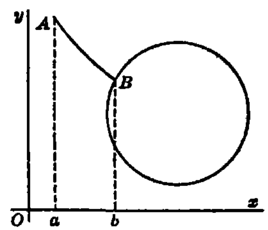


FIG. 108

This example is typical for the general case of the integral

$$(1) \qquad J = \int_a^b F(x, y, y')\, dx,$$

where all the curves $y = f(x)$ admitted to competition start from one and the same fixed point, A, but end at any point of a given curve

$$\Gamma: \qquad y = \phi(x),$$

where $\phi(x)$ is continuous, together with its first derivative.

First of all, it is clear that a necessary condition for the function y which makes the integral a minimum (or a maximum) is that it be an *extremal*, *i.e.* a solution of Euler's Equation. For, an admissible set of varied curves is that for which both end points are fixed, and this is the case of § 2.

Next, let a pencil of varied curves $\bar{C}$ be chosen as follows. Suppose the extremal

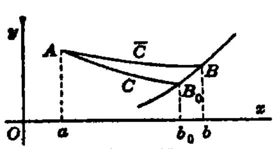


FIG. 109

$$C: \qquad y = f(x), \qquad a \leqq x \leqq b_0,$$

to be the solution of the problem, and suppose C can be extended slightly beyond B_0. Thus $f(x)$ will be continuous, together with its first derivative, throughout a somewhat larger interval, $a \leqq x \leqq b'$, where $b' > b_0$, and will still satisfy Euler's Equation.

Let $\eta = \eta(x)$ be chosen continuous, together with its first derivative, in the above interval, and let

$$(2) \qquad \eta(a) = 0, \qquad \eta(b_0) \neq 0.$$

Then $\bar{C}$ shall be defined by the equation:

$$\bar{C}: \qquad \bar{y} = y + k\eta,$$

where $y = f(x)$ and where k is so chosen that $\overline{C}$ goes through B, or

(3) $$f(b) + k\eta(b) = \phi(b).$$

Let $b = b_0 + \epsilon$. Then $k = k(\epsilon)$ is given by the equation:

(4) $$k(\epsilon) = \frac{\phi(b_0 + \epsilon) - f(b_0 + \epsilon)}{\eta(b_0 + \epsilon)}.$$

Since $\phi(b_0) = f(b_0)$, it follows that $k(0) = 0$ and

(5) $$k'(0) = \frac{\phi'(b_0) - f'(b_0)}{\eta(b_0)} = \frac{\tan\omega - y'}{\eta}\bigg|_{b=b_0},$$

where ω denotes the angle from the axis of x to the tangent to Γ.

Consider the integral J formed for the varied curve $\bar{y}$. It is a function of ϵ:

$$J(\epsilon) = \int_a^b F(x, y + k\eta, y' + k\eta')\, dx,$$

and it has a minimum (or maximum) when $\epsilon = 0$. Hence

$$(dJ/d\epsilon)_{\epsilon=0} = 0.$$

Now

$$\left(\frac{dJ}{d\epsilon}\right)_{\epsilon=0} = \int_a^{b_0} \left\{k'(0)\,\eta F_y + k'(0)\,\eta' F_{y'}\right\} dx + F\bigg|_{x=b_0}.$$

The integral can be transformed by integration by parts, as in § 2:

$$\int_a^{b_0} \eta' F_{y'}\, dx = \eta F_{y'}\bigg|_a^{b_0} - \int_a^{b_0} \eta \frac{dF_{y'}}{dx}\, dx.$$

Thus

$$\left(\frac{dJ}{d\epsilon}\right)_{\epsilon=0} = k'(0) \int_a^{b_0} \eta\left(F_y - \frac{dF_{y'}}{dx}\right) dx + \left\{k'(0)\,\eta F_{y'} + F\right\}_{x=b_0} = 0.$$

This latter integral vanishes, since y is an extremal. On substituting for $k'(0)$ its value from (5), we have:

$$(\tan\omega - y')\, F_{y'} + F = 0,$$

or

(6) $$F_{y'} \sin\omega + (F - y'F_{y'}) \cos\omega = 0,$$

where the functions F, $F_{y'}$ are formed for the point $x = b_0$, $y = f(b_0)$, $y' = f'(b_0)$. This is the condition we set out to obtain.

Similarly, the right-hand end point may be fixed and the left-hand one variable. In that case, a condition of the same form as (6) holds for the left-hand end point. Or, finally, both end points may be variable, and we then have the two equations (6) holding independently.

The Catenary. Returning now to the example with which the paragraph began, we have

$$F(x, y, y') = y\sqrt{1 + y'^2}.$$

Thus (6) gives here:

$$\tan\omega = -\frac{1}{y'},$$

or the catenary meets the circle at right angles.

· *The Isoperimetric Case,*

$$J = \int_a^b F(x, y, y')\,dx, \qquad K = \int_a^b \Phi(x, y, y')\,dx.$$

Here we introduce the varied curve

$\bar{C}$: $$\bar{y} = y + k\eta + l\zeta$$

and determine k and l from the equations

$$K = \int_a^b \Phi(x, y + k\eta + l\zeta,\ y' + k\eta' + l\zeta')\,dx,$$

$$\phi(b_0 + \epsilon) = f(b_0 + \epsilon) + k\,\eta(b_0 + \epsilon) + l\,\zeta(b_0 + \epsilon).$$

These equations can be solved for k and l, provided the Jacobian

$$\left|\begin{matrix} \dfrac{\partial K}{\partial k} & \dfrac{\partial K}{\partial l} \\ \eta & \zeta \end{matrix}\right|_{\epsilon=0} \neq 0.$$

The functions η and ζ can be chosen so that this condition will be fulfilled unless

$$\int_a^b (\eta\Phi_y + \eta'\Phi_{y'})\,dx$$

vanishes for all choices of η. This difficulty does not occur in the cases which arise in practice.

Instead next of setting $(\partial J/\partial\epsilon)_{\epsilon=0} = 0$, it will be more convenient to consider the integral

$$\int_c^b \left\{F(x, \bar{y}, \bar{y}') + \lambda\Phi(x, \bar{y}, \bar{y}')\right\}dx,$$

which is merely $J+$ const. On requiring that its derivative with respect to ϵ vanish for $\epsilon=0$, we are led first to the condition

$$\int_a^{h_0}\left\{(k'\eta+l'\zeta)\overline{F_y+\lambda\Phi_y}+(k'\eta'+l'\zeta')\overline{F_{y'}+\lambda\Phi_{y'}}\right\}dx+\overline{F+\lambda\Phi}\Big|_{x=h_0}=0.$$

From this point on the analysis is similar to that of the earlier case, and we find as the analogue of (6) the equation

$$(7)\qquad \overline{F_{y'}+\lambda\Phi_{y'}}\sin\omega+\left\{\overline{F+\lambda\Phi}-y'\overline{F_{y'}+\lambda\Phi_{y'}}\right\}\cos\omega=0.$$

EXERCISES

1. A uniform flexible heavy string has its ends fastened to weightless rings which slide on smooth fixed wires in a vertical plane. Show that, when the string is in equilibrium, it meets the wires at right angles.

2. A variable curve C of given length connects two fixed curves, C_1 and C_2, which lie above the axis of x. Show that, when the area bounded by C, the two ordinates at its extremities, and the axis of x is least, C does not in general meet C_1 or C_2 at right angles.

3. If, in the preceding question, the area bounded by C, C_1 and C_2 (where now C_1 and C_2 are supposed to meet under C) is to be made a minimum, then C will meet C_1 and C_2 at right angles.

10. Parametric Form and the so-called "Variation of the Independent Variable." In the integral

$$(1)\qquad J=\int_a^b F\left(x,y,\frac{dy}{dx}\right)dx$$

let the curve $y=f(x)$ be represented parametrically:

$$(2)\qquad x=\phi(\tau),\qquad y=\psi(\tau),\qquad \tau_0\leqq\tau\leqq\tau_1,$$

where ϕ and ψ are continuous, together with their first derivatives, and the latter do not vanish simultaneously. Let

$$(3)\qquad \frac{dx}{d\tau}=x'_\tau,\qquad \frac{dy}{d\tau}=y'_\tau,\qquad \frac{dy}{dx}=y'.$$

Then

$$(4)\qquad J=\int_{\tau_0}^{\tau_1}\Phi(x,y,x'_\tau,y'_\tau)\,d\tau,$$

where

$$\Phi(x, y, x'_\tau, y'_\tau) = F\left(x, y, \frac{y'_\tau}{x'_\tau}\right)x'_\tau.$$

Euler's Equation here takes on the form:

(5) $$\Phi_x - \frac{d}{d\tau}\Phi_{x'_\tau} = 0, \qquad \Phi_y - \frac{d}{d\tau}\Phi_{y'_\tau} = 0.$$

Since

$$\Phi_y - \frac{d}{d\tau}\Phi_{y'_\tau} = F_y x'_\tau - \frac{d}{d\tau}F_{y'} = x'_\tau\left(F_y - \frac{d}{dx}F_{y'}\right),$$

it is seen that the second of these equations is equivalent to Euler's Equation in the earlier form,

(6) $$F_y - \frac{d}{dx}F_{y'} = 0,$$

provided $\phi'(\tau) \neq 0$.

Let x and y receive the variations ξ and η respectively,* where, however, we no longer demand that ξ, η vanish at the extremities of the interval, (τ_0, τ_1). Then

(7) $$\delta J = \int_{\tau_0}^{\tau_1} (\xi\Phi_x + \eta\Phi_y + \xi'_\tau\Phi_{x'_\tau} + \eta'_\tau\Phi_{y'_\tau})\, d\tau.$$

In terms of F the integral has the value:

(8) $$\delta J = \int_{\tau_0}^{\tau_1} (\xi x'_\tau F_x + \eta x'_\tau F_y + \xi'_\tau[F - \frac{y'_\tau}{x'_\tau}F_{y'}] + \eta'_\tau F_{y'})\, d\tau.$$

If, in particular, $\phi'(\tau) > 0$ throughout the interval, it is possible to change the variable of integration from τ to x, and thus

(9) $$\delta J = \int_a^b \left\{ \xi F_x + \xi'[F - y'F_{y'}] + \eta F_y + \eta' F_{y'} \right\} dx$$

* The varied curve,

$$X = x + \xi = \Phi(\tau), \qquad Y = y + \eta = \Psi(\tau), \qquad \tau_0 \leqq \tau \leqq \tau_1,$$

is known as a *strong variation*, since its slope no longer necessarily differs but slightly from that of the original curve at corresponding points. Indeed, Y is not necessarily a single-valued function of X. But ξ, η, ξ'_τ, η'_τ, are assumed to be continuous and numerically small throughout the interval, and thus $\Phi'(\tau)$ and $\Psi'(\tau)$ will not vanish simultaneously. In distinction, the varied curve of § 1 is called a *weak variation*.

This integral can be transformed by integration by parts, as was done in a similar case in § 2, with the result

$$(10)\qquad \delta J=\int_a^b\left\{\xi\left(F_x-\frac{d}{dx}[F-y'F_{y'}]\right)+\eta\left(F_y-\frac{dF_{y'}}{dx}\right)\right\}dx$$
$$+\left\{\xi[F-y'F_{y'}]+\eta F_{y'}\right\}_{x=a}^{x=b}.$$

Since

$$\frac{d}{dx}[F-y'F_{y'}]=F_x+y'F_y-y'\frac{dF_{y'}}{dx},$$

we can give to δJ the form:

$$(11)\qquad \delta J=\int_a^b(\delta y-y'\delta x)\left(F_y-\frac{dF_{y'}}{dx}\right)dx$$
$$+\left\{[F-y'F_{y'}]\,\delta x+F_{y'}\,\delta y\right\}_{x=a}^{x=b},$$

where $\delta x=\xi$ and $\delta y=\eta$.

If ξ and η are required to vanish at the end points of the interval the last term drops out. Formula (11) still holds true, even when this condition is not imposed, provided δx and δy are suitably defined; but a more elaborate definition of these variations is there needed.*

"*Variation of the Independent Variable.*" We can, in particular, allow the function $\phi(\tau)$ to be the function τ. Then the first of equations (2) becomes $x=\tau$. The variation, ξ, of this function is, however, no more and no less general than before, and since $\phi'(\tau)>0$ here, we arrive at the same final result, namely, equation (11). Because the dependent variable x is here equal to the independent variable τ, this case is sometimes described as the "variation of the independent variable." But the expression is misleading

* Cf. the admirable treatment of this point in Bolza, *Vorlesungen über Variationsrechnung*, p. 45. One of the weak points in the use of the Calculus of Variations in physics lies in the tacit assumption that the variations require no particular definition, for everyone knows what δy, δJ, etc., mean. As a matter of fact, their definition, except in the simplest cases like those of §§ 2, 6, is an exceedingly delicate matter, and Bolza's contribution is most valuable.

Even in so simple a matter as the isoperimetric problem of § 8 the foregoing definitions of δy, $\delta y'$ (§ 5), will not lead to the equation $\delta K=0$, for this equation will not be true; but suitably modified definitions of δy, $\delta y'$ will lead to the equation $\delta K=0$ as a necessary condition.

if taken literally, for it is a contradiction in terms, — we do not define "variation" for other than dependent variables. The language is of a piece with that used to describe the method in differential equations known as the "variation of constants", and with the expression "an infinitesimal constant." It is little short of an Irish bull.

EXERCISES

1. Extend the method to the isoperimetric case,

$$J=\int_a^b F(x, y, y')\,dx, \qquad K=\int_a^b \Phi(x, y, y')\,dx$$

and show that here equation (11) becomes

$$(12)\quad \delta J=\int_a^b(\delta y-y'\delta x)\left\{\overline{F_y+\lambda\Phi_y}-\frac{d}{dx}\overline{F_{y'}+\lambda\Phi_{y'}}\right\}dx+$$

$$\left\{[\overline{F+\lambda\Phi}-y'\,\overline{F_{y'}+\lambda\Phi_{y'}}]\,\delta x+\overline{F_{y'}+\lambda\Phi_{y'}}\,\delta y\right\}_{x=a}^{x=b}.$$

2. Extend the method to multiple integrals and show that, for the double integral

$$J=\iint_S F\left(x, y, u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}\right)dS,$$

equation (11) becomes

$$(13)\quad \delta J=\iint_S(\delta u-u_x\delta x-u_y\delta y)\left(F_u-\frac{\partial}{\partial x}F_{u_x}-\frac{\partial}{\partial y}F_{u_y}\right)dS,$$

provided that the variations $\delta x=\xi$, $\delta y=\eta$, $\delta u=\zeta$ all vanish identically on the boundary. Otherwise, a line integral must be added, which is the counterpart of the last term in (11).

Suggestion. Develop first the formulas which correspond to (4), (7), and (8), when

$$x=\phi(\lambda, \mu), \qquad y=\psi(\lambda, \mu),$$

remembering that (cf. p. 150, Ex. 31)

$$\frac{\partial u}{\partial x}=\frac{\partial(u, y)}{\partial(\lambda, \mu)}\Big/\frac{\partial(x, y)}{\partial(\lambda, \mu)}, \qquad \frac{\partial u}{\partial y}=\frac{\partial(x, u)}{\partial(\lambda, \mu)}\Big/\frac{\partial(x, y)}{\partial(\lambda, \mu)}.$$

3. Work Ex. 2 for the volume integral

$$J=\iiint_V F\left(x, y, z, u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial u}{\partial z}\right) dV,$$

writing down in advance by inspection the probable answer.

11. Hamilton's Principle. Consider a single particle, of mass m, situated at any point $P:(x, y, z)$ and acted on by the attraction of a particle situated at the origin. The magnitude of the force, measured in the absolute units of dynamics, will be $m\lambda/r^2$, and its components,

$$(1) \qquad X=-m\lambda\frac{x}{r^3}, \qquad Y=-m\lambda\frac{y}{r^3}, \qquad Z=-m\lambda\frac{z}{r^3}.$$

There exists a *force function* (p. 146):

$$(2) \qquad U=\frac{m\lambda}{r},$$

whose derivative in any direction gives the component of the force in that direction. In particular, it is seen that

$$(3) \qquad \frac{\partial U}{\partial x}=X, \qquad \frac{\partial U}{\partial y}=Y, \qquad \frac{\partial U}{\partial z}=Z.$$

The motion of m under the influence of the force that acts is determined by Newton's Second Law of Motion, *Introduction to the Calculus*, Chap. XIII, § 1:

$$(4) \qquad m\frac{d^2x}{dt^2}=X, \qquad m\frac{d^2y}{dt^2}=Y, \qquad m\frac{d^2z}{dt^2}=Z.$$

Its path is represented most naturally in parametric form, the time t being the parameter:

$$(5) \qquad x=f(t), \qquad y=\phi(t), \qquad z=\psi(t).$$

Hamilton's Integral. Consider the integral:

$$(6) \qquad I=\int_{t_0}^{t_1}(T+U)\,dt,$$

where

$$T=\tfrac{1}{2}m(\dot{x}^2+\dot{y}^2+\dot{z}^2), \qquad \dot{x}=\frac{dx}{dt}, \text{ etc.}$$

denotes the kinetic energy, and U, the force function (2), is the negative of the potential energy. This integral has an altogether definite value for the actual motion of the particle under the action of the given force. We can, however (by applying suitable extraneous forces), cause the particle to go from its first position at (x_0, y_0, z_0) to its second position at (x_1, y_1, z_1) along a different path, and we can choose arbitrarily the law of the velocity it shall have along this path. We will, however, agree that the total time from the first position to the second shall be the same as in the case of the natural motion. The integral I will have a definite value for this second path, and it will be positive, since the integrand is positive.

The totality of such values of the integral must have a lower limit, and it is not unlikely that, for some path, this lower limit will actually be realized, and thus the integral will have a minimum. A necessary condition that this be the case is that Euler's Equation be satisfied; *i.e.* that

$$\frac{\partial(T+U)}{\partial x} - \frac{d}{dt}\frac{\partial(T+U)}{\partial \dot{x}} = 0, \tag{7}$$

or

$$m\frac{d^2x}{dt^2} = X,$$

with two similar equations for y and z. But these three equations are precisely the equations (4) which govern the free motion, with no extraneous forces acting.

We may say, therefore, assuming that there is one and only one path which makes Hamilton's Integral a minimum, that this is the path the particle will follow when acted on only by the forces given at the outset, and we are thus led to

HAMILTON'S PRINCIPLE. *The path which the particle follows when acted on by the given forces is the path that makes Hamilton's Integral a minimum.*

The case of a system of n particles, $m_i : (x_i, y_i, z_i)$, attracted by ν fixed particles under the law of universal gravitation admits a precisely similar treatment. Here

$$T = \sum_i \frac{m_i}{2}(\dot{x}_i^2 + \dot{y}_i^2 + \dot{z}_i^2), \qquad U = \sum_{i,j}\frac{m_i \lambda_j}{r_{ij}}, \tag{8}$$

$$r_{ij}^2 = (x_i - a_j)^2 + (y_i - b_j)^2 + (z_i - c_j)^2.$$

The integral (6) has the same form as before, and the $3n$ equations of Euler are precisely the $3n$ equations given by Newton's Second Law of Motion, applied to each of the n particles, namely:

$$(9)\quad m_i\frac{d^2x_i}{dt^2}=\frac{\partial U}{\partial x_i},\qquad m_i\frac{d^2y_i}{dt^2}=\frac{\partial U}{\partial y_i},\qquad m_i\frac{d^2z_i}{dt^2}=\frac{\partial U}{\partial z_i},\quad i=1,2,\cdots,n.$$

Thus in this case we have *proved* Hamilton's Principle *mathematically* by means of Newton's Second Law, assuming that Hamilton's Integral has a minimum. And now we extend that Principle to cases not covered by other physical laws and make that Principle the defining element, the new physical postulate.* But the above form of statement, namely, as a minimum principle, is not to be retained, as is shown by Example 2 below. It is rather the *stationary* principle, which consists in requiring that the variation of the integral (6) vanish, or $\delta I=0$.

GENERALIZED FORM OF HAMILTON'S PRINCIPLE. *Let the kinetic energy of a material system be denoted by T, and its potential energy by $-U$. Then its motion is such that the variation of Hamilton's Integral vanishes:*

$$\delta\int_{t_0}^{t_1}(T+U)\,dt=0.$$

Here, t_0 and t_1 are any two values of the time, t, which is the variable of integration, and hence the independent variable of the function $T+U$. Only such variations are considered as leave the end points fixed; i.e. t_0 and t_1 are fixed, and all the dependent variables (but not their derivatives) must, for all variations, preserve the same value when $t=t_0$, and their values for $t=t_1$ must likewise be preserved.†

Example 1. Consider a system of particles having n degrees of freedom, but subjected to smooth constraints that are not moving. Let the natural coordinates of the system be $q_1,\cdots,q_n$, and let the

* A "Principle" in Mechanics has been well described by Mr. B. O. Koopman as follows: "According to the usage of the present day the word *principle* in physics has lost its metaphysical implication, and now denotes a physical truth of a certain importance and generality. Like all physical truths, it rests ultimately on experiment; but whether it is taken as a physical law, or appears as a consequence of physical laws already laid down, does not matter."

† We add the remark, which would be superfluous except for great confusion shown in the literature, that in Hamilton's Principle the time can not be varied. Of course it cannot, for to "vary the independent variable" is to introduce a contradiction in terms, to do violence to the definition of a *variation*; *cf.* § 10.

kinetic energy be

$$T=\sum A_{ij}\dot{q}_i\dot{q}_j, \qquad i, j=1, \cdots, n,$$

where A_{ij} is a function of $q_1, \cdots, q_n$. Let the forces form a conservative system, and let the potential energy under these forces be a function of $q_1, \cdots, q_n$. Let the negative of the potential energy be denoted by U. Then a necessary and sufficient condition for the motion of the system is given by the vanishing of the variation of Hamilton's Integral:

$$\delta\int_{t_0}^{t_1}(T+U)\,dt=0.$$

For, Euler's Equations here become

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_i}+\frac{\partial(T+U)}{\partial q_i}=0, \qquad i=1, \cdots, n,$$

and these are precisely Lagrange's Equations.

Example 2. To find the differential equation of the vibrating string. Consider, for example, the motion of a piano string or a violin string, the ends of which are fixed. Let the motion take place in a fixed plane, and assume (*i*) that no point of the string moves far from its position when the whole string is at rest; and (*ii*) that the greatest angle which the string makes at any point with its line of equilibrium is small.

Approximations are now introduced as follows. (a) The component of the motion of any point parallel to the axis of x (Fig. 96) is neglected, and thus the kinetic energy is given by the integral

$$T=\int_0^l\frac{\rho}{2}\left(\frac{\partial y}{\partial t}\right)^2dx,$$

where ρ denotes the density of the string, assumed constant.

(b) The potential energy is proportional to the square of the stretching, assumed uniform: $s^2=(l'-l_0)^2$. Let $\sigma=l'-l$. Then, $l'-l_0=\sigma+(l-l_0)$, and thus the potential energy, diminished by a constant, is approximately proportional to σ, since σ^2 can be neglected. Now

$$\sigma=\int_0^l\sqrt{1+\left(\frac{\partial y}{\partial x}\right)^2}\,dx-l.$$

Since $\partial y/\partial x$ is small by hypothesis, the radical is seen to be approximately equal to $1+\frac{1}{2}(\partial y/\partial x)^2$, and hence U, which is the negative of the potential energy, is given approximately by the expression:

$$U = \int_0^l -\frac{p}{2}\left(\frac{\partial y}{\partial x}\right)^2 dx,$$

where p is a constant.

We thus arrive at the following form for Hamilton's Integral, corresponding to the above approximations:

$$\int_{t_0}^{t_1}(T + U)\,dt = \int_{t_0}^{t_1}\int_0^l \left[\frac{\rho}{2}\left(\frac{\partial y}{\partial t}\right)^2 - \frac{p}{2}\left(\frac{\partial y}{\partial x}\right)^2\right] dx\,dt.$$

This is equivalent to the double integral of the bracket, extended over the fixed rectangle $t_0 \leqq t \leqq t_1$; $0 \leqq x \leqq l$. And now Hamilton's Principle consists in setting the variation of this integral equal to zero. From this condition follows, as a necessary condition, Euler's Equation, and thus we have:

$$\frac{\partial}{\partial t}\left(\rho\,\frac{\partial y}{\partial t}\right) - \frac{\partial}{\partial x}\left(p\,\frac{\partial y}{\partial x}\right) = 0,$$

or

$$\frac{\partial^2 y}{\partial t^2} = a^2\,\frac{\partial^2 y}{\partial x^2}, \qquad a^2 = p/\rho,$$

as the differential equation of the vibrating string.

It would be a mistake, however, to think that Hamilton's Integral attains its minimum value when y is a solution of Euler's Equation. Consider the case $\rho = p = 2$, $a = 1$, $l = \pi$. Then

$$I = \int_{t_0}^{t_1}\int_0^{\pi}\left[\left(\frac{\partial y}{\partial t}\right)^2 - \left(\frac{\partial y}{\partial x}\right)^2\right] dx\,dt; \qquad \frac{\partial^2 y}{\partial t^2} = \frac{\partial^2 y}{\partial x^2}.$$

A solution of Euler's Equation is here

$$y = \sin t \sin x.$$

Let y be varied as follows:

$$Y = y + \zeta, \qquad \zeta = h \sin\frac{n\pi(t - t_0)}{t_1 - t_0}\sin mx,$$

where h is a constant which may be chosen arbitrarily small, and n and m are integers. Then

$$\int_{t_0}^{t_1}\int_0^{\pi}\left[\left(\frac{\partial(y+\zeta)}{\partial t}\right)^2-\left(\frac{\partial(y+\zeta)}{\partial x}\right)^2\right]dx\,dt=\int_{t_0}^{t_1}\int_0^{\pi}\left[\left(\frac{\partial y}{\partial t}\right)^2-\left(\frac{\partial y}{\partial x}\right)^2\right]dx\,dt$$

$$+\int_{t_0}^{t_1}\int_0^{\pi}\left(2\zeta_t\frac{\partial y}{\partial t}-2\zeta_x\frac{\partial y}{\partial x}\right)dx\,dt+\int_{t_0}^{t_1}\int_0^{\pi}\left[\left(\frac{\partial \zeta}{\partial t}\right)^2-\left(\frac{\partial \zeta}{\partial x}\right)^2\right]dx\,dt.$$

The first integral in the last line vanishes because it is the variation of I, formed for the extremal y. And now the last integral can be computed directly and is seen to have the value:

$$\frac{\pi h^2}{4}\left\{\frac{\pi^2 n^2}{t_1-t_0}-m^2(t_1-t_0)\right\}.$$

By giving m and n suitable values, we can make the brace change sign.

Thus Hamilton's Integral is in this case neither a maximum nor a minimum for the extremal y; but it is none the less *stationary*, for its variation vanishes.

Example 2 illustrates a further point. Hamilton's Principle as stated above applies to the simple integral (6), and it is the variation of this integral that is to vanish. But in solving the problem of Example 2 we have set the variation of a wholly different integral — the double integral — equal to 0. Clearly, it must be shown that these two conditions are equivalent. It is not difficult to do so in this case, and the student will do well to work out the proof. The incident brings out clearly the fact that Hamilton's Principle depends for its very statement on the definition of *variation* (δ), and so the formulation of this definition must precede any application of the Principle. This question is treated by Bolza; cf. the reference given above in § 10.

EXERCISES

1. If the string is allowed to vibrate in three dimensions, show that its motion is governed by the simultaneous equations:

$$\frac{\partial^2 y}{\partial t^2}=a^2\frac{\partial^2 y}{\partial x^2},\qquad \frac{\partial^2 z}{\partial t^2}=a^2\frac{\partial^2 z}{\partial x^2}.$$

2. Let the string vibrate longitudinally, *i.e.* in its own line, and let a point which, at rest, had the coordinate x have the coordinate $x_1=x+u$. Show that

$$\frac{\partial^2 u}{\partial t^2}=b^2\frac{\partial^2 u}{\partial x^2}.$$

The most general motion of the string is a superposition of the motions of Ex. 1 and 2.

3. *Vibrating Membrane.* Consider a vibrating membrane, like a drum-head, whose position of equilibrium is in the (x, y)-plane. If its points move approximately in right lines orthogonal to the plane; if, moreover, their excursions are small and the tangent plane to the distorted surface makes at most a small angle with the (x, y)-plane, then the potential energy is approximately

$$\frac{p}{2}\int\int_S \left\{\left(\frac{\partial z}{\partial x}\right)^2+\left(\frac{\partial z}{\partial y}\right)^2\right\} dS.$$

Find the kinetic energy, T, and show that the differential equation which governs the motion is

$$\frac{\partial^2 z}{\partial t^2}=a^2\left\{\frac{\partial^2 z}{\partial x^2}+\frac{\partial^2 z}{\partial y^2}\right\}.$$

4. *Vibrating Rod.* Let a uniform straight rod vibrate in a fixed plane. It is assumed not only that the displacements of the points of the rod are small, but also that the angle which the rod makes at any point with its position of rest is so small that the contribution to the potential energy due to the change in length of the rod is negligible. It can then be shown that the potential energy of the rod is proportional to the integral of the square of the curvature Show that Hamilton's Integral here becomes

$$\int_{t_0}^{t_1}\int_0^l\left\{\frac{\rho}{2}\left(\frac{\partial u}{\partial t}\right)^2-\frac{c}{2}\left(\frac{\partial^2 u}{\partial x^2}\right)^2\right\} dx\, dt,$$

and hence Euler's Equation takes the form

$$\frac{\partial^2 u}{\partial t^2}+a^2\frac{\partial^4 u}{\partial x^4}=0, \qquad a^2=\frac{c}{\rho}.$$

5. *Vibrating Plate.* Let a uniform plane plate vibrate transversely. It is assumed that the normal at any point makes so small an angle with the plane of rest, the (x, y)-plane, that the contribution to the potential energy due to the inclination of the normal is negligible. The equation of the plate, in the neighborhood of any point (x_0, y_0, u_0), is then approximately

$$u=ax'^2+2hx'y'+by'^2, \qquad x=x_0+x', \qquad y=y_0+y'.$$

The potential energy of the plate can be shown to be the double integral of a homogeneous quadratic expression in the three coefficients

$$a = \frac{1}{2}\frac{\partial^2 u}{\partial x^2}, \qquad h = \frac{1}{2}\frac{\partial^2 u}{\partial x \partial y}, \qquad b = \frac{1}{2}\frac{\partial^2 u}{\partial y^2}.$$

Since this expression must clearly be invariant of a rotation of the axes, it must depend on the two invariants of the above quadratic form, namely,

$$\Omega = a + b = \frac{1}{2\rho_1} + \frac{1}{2\rho_2} = \frac{1}{2} \text{ (mean curvature)}$$

$$\Theta = ab - h^2 = \frac{1}{\rho_1 \rho_2} = \text{ total curvature.}$$

Hence the expression must be of the form

$$\lambda \Omega^2 + \mu \Theta,$$

where λ and μ are physical constants depending on the plate. We have, then, finally for the potential energy the expression

$$\iint_S \left\{ \frac{\lambda}{2}\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right)^2 + \frac{\mu}{2}\left[\left(\frac{\partial^2 u}{\partial x \partial y}\right)^2 - \frac{\partial^2 u}{\partial x^2}\frac{\partial^2 u}{\partial y^2}\right]\right\} dS.$$

Show that the differential equation of the vibrating plate is

$$\rho \frac{\partial^2 u}{\partial t^2} + \lambda\left(\frac{\partial^4 u}{\partial x^4} + 2\frac{\partial^4 u}{\partial x^2 \partial y^2} + \frac{\partial^4 u}{\partial y^4}\right) = 0.$$

6. *Propagation of Sound.* In the case of sound waves it can be shown that the potential energy of the medium (air, water, iron, etc.) in which the disturbance takes place is proportional approximately to

$$\iiint_V \left(\frac{\partial u}{\partial t}\right)^2 dV,$$

where u denotes the velocity potential which governs the motion of the individual particles (more properly, of the points of the material distribution). The kinetic energy is proportional approximately to

$$\iiint_V \left\{\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial u}{\partial z}\right)^2\right\} dV.$$

Show that the partial differential equation which governs the phenomenon of sound waves is

$$\frac{\partial^2 u}{\partial t^2} = c^2 \left\{\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}\right\},$$

where u denotes the velocity potential.

12. Least Action. Let us consider, with Jacobi, the following integral, known as the *action*:

$$(1) \qquad \mathfrak{A} = \int_{t_0}^{t_1} \sqrt{2(U+h)}\,\sqrt{T}\,dt,$$

where, as before, T and U denote respectively the kinetic energy and the negative of the potential energy, and h is a constant so chosen that, in the actual motion, *i.e.* the motion of the system under the action of the given forces,

$$(2) \qquad T = U + h.$$

We impose now, however, a further restriction, namely, that *neither T nor U shall depend explicitly on t.* It is assumed that U depends only on the coordinates of the system, and not on their time derivatives. The integral thus ceases to depend on t, — at least in those cases in which T is a homogeneous quadratic form in the time derivatives of the coordinates of the system, as, for example, when

$$(3) \qquad T = \sum_i \frac{m_i}{2}[\dot{x}_i^2 + \dot{y}_i^2 + \dot{z}_i^2],$$

or when T, in terms of the so-called *natural coordinates*, $q_1, q_2, \cdots, q_\lambda$, of the system, has the form

$$(4) \qquad T = \sum_{i,j} A_{ij}\dot{q}_i\dot{q}_j,$$

where A_{ij} depends on $q_1, \cdots, q_\lambda$, but not on their time derivatives.

In these cases we may eliminate t even in form by introducing a parameter, τ, which lies in a fixed interval, $\tau_0 \leqq \tau \leqq \tau_1$, and setting

$$(5) \qquad t = \phi(\tau),$$

where $\phi'(\tau)$ and $\phi''(\tau)$ are continuous in the above interval, and $\phi'(\tau) > 0$ at all points of the interval.

Thus the motion of the system is described completely in terms of the parameter τ. Let

$$(6) \qquad S = \sum_i \frac{m_i}{2}[x_i'^2 + y_i'^2 + z_i'^2], \qquad x_i' = \frac{dx_i}{d\tau}, \quad \text{etc.}$$

Then

$$(7) \qquad S = T\,\frac{dt^2}{d\tau^2},$$

and the action, $\mathfrak{A}$, assumes the form:

$$(8)\qquad \mathfrak{A} = \int_{\tau_0}^{\tau_1} \sqrt{2(U+h)}\,\sqrt{S}\,d\tau.$$

We now proceed to vary the path, holding the end points (x_i^0, y_i^0, z_i^0) and (x_i^1, y_i^1, z_i^1) fast. Through the application of suitable extraneous forces we can cause the system to describe the varied path at any rate we choose. And now we make use of this freedom of choice to demand that *the rate at which the system describes the varied path shall be such that, at each instant,*

$$(9)\qquad T = U + h$$

along the varied path.

This last condition has as a consequence that the time will not in general be the same for the varied path. Why should it be?

Since the force function, U, depends only on the position of the individual particles, T is determined by (9) as a function of τ. But S is also determined by definition as a function of τ. Hence t is determined as a function of τ by (7), or

$$(10)\qquad t = \int_{\tau_0}^{\tau} \sqrt{\frac{S}{U+h}}\,d\tau.$$

The Minimum Principle. Among all possible paths, it is reasonable to inquire what path will make the action a minimum. For, the integrand being always positive, we see that the action is positive, and so it must have a lower limit. Suppose, then, that this lower limit is reached, *i.e.* that there is a path for which the integral is a minimum, and suppose the requisite derivatives exist and are continuous for this path. Then Euler's Equation will be satisfied, and, in the case of n particles moving without constraint, we shall have:

$$\frac{\partial}{\partial x_i}(\sqrt{2(U+h)}\sqrt{S}) - \frac{d}{d\tau}\frac{\partial}{\partial x_i'}(\sqrt{2(U+h)}\,\sqrt{S}) = 0,$$

or

$$(11)\qquad \frac{\sqrt{S}}{\sqrt{U+h}}\frac{\partial U}{\partial x_i} - \frac{d}{d\tau}\left(\frac{\sqrt{U+h}}{\sqrt{S}}\,m_i x_i'\right) = 0, \quad i = 1, 2, \cdots, n,$$

with corresponding equations in y_i and z_i.

Thus we have $3n$ differential equations for determining the $3n$ coordinates x_i, y_i, z_i of the system as functions of the arbitrary parameter τ. The time has been completely eliminated from the problem of determining the path, and now the time is subsequently determined by (10).

The choice of the parameter τ is arbitrary, subject merely to the conditions of continuity (including derivatives) stated above. We may, then, in the final equation (11), choose τ as $=t$, *i.e.* set $\phi(\tau)=\tau$. Then (11) reduces to

$$(12) \qquad m_i\frac{d^2x_i}{dt^2}=\frac{\partial U}{\partial x_i},$$

and the $3n$ equations thus arising are precisely the $3n$ equations which express Newton's Second Law of Motion for the particles of the given system.

The result may, however, be obtained without specializing the choice of τ. For, if $d\tau$, as given by (10), be substituted in (11), the latter equations take on the form (12).

Thus we have proved the Principle of Least Action in the case of n particles acted on by forces which admit a force function, on the assumption that the action integral, $\mathfrak{A}$, has a minimum. As in the case of Hamilton's Principle, so here, the minimum principle demands more than is needed. It is enough to require that the integral be *stationary*. In fact, there are very simple cases in which the action is not a minimum; but its variation vanishes. For example, let a particle be projected vertically upward, and let z denote its distance below the highest point of its path. Then

$$T=\frac{m}{2}\left(\frac{dz}{dt}\right)^2, \qquad S=\frac{m}{2}\left(\frac{dz}{d\tau}\right)^2, \qquad U+h=mgz,$$

$$\mathfrak{A}=m\sqrt{g}\int_{\tau_0}^{\tau_1}\sqrt{z}\sqrt{\left(\frac{dz}{d\tau}\right)^2}\,d\tau.$$

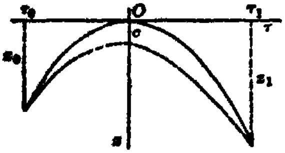


FIG. 110

Its natural path is given by the equations:

$$z=\tfrac{1}{2}g\tau^2, \qquad t=\tau, \qquad \tau_0\leqq\tau\leqq\tau_1.$$

Consider now a varied path having the same end points (τ_0, z_0) and (τ_1, z_1), and having

$$\left(\frac{dz}{d\tau}\right)_{\tau=0}=0, \qquad z\,|_{\tau=0}=c>0.$$

The value of the action for this path is seen to be:

$$\mathfrak{A} = \tfrac{2}{3} m \sqrt{g}\left(z_0^{\frac{3}{2}} + z_1^{\frac{3}{2}} - 2c^{\frac{3}{2}}\right),$$

and the time is:

$$t_1 - t_0 = \sqrt{\frac{2}{g}}\left(\sqrt{z_0} + \sqrt{z_1} - 2\sqrt{c}\right).$$

The action and the time for the natural path are obtained from these formulas by putting $c = 0$. Thus the action is *less* for the varied path, — and the time is also less.

In the foregoing example, Euler's Equation becomes an identity, except for $z = 0$, when it ceases to have a meaning.

Elimination of the Time. The principle of least action eliminates the time from the problem. The motion is determined in terms of τ, and then the time is computed by a quadrature:

$$T\left(\frac{dt}{d\tau}\right)^2 = S, \qquad \left(\frac{dt}{d\tau}\right)^2 = \frac{S}{U+h}, \qquad t = \int_0^\tau \sqrt{\frac{S}{U+h}}\, d\tau.$$

This last equation may be looked on as the *definition* of the time.

Example 1. To determine the path of a projectile, the resistance of the air being neglected.

Let the axis of x be chosen vertical and positive downwards. Then

$$U = mg\,x,$$

provided the level of the (y, z)-plane is suitably determined Furthermore,

$$T = \frac{m}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2).$$

Hence

$$\mathfrak{A} = m\sqrt{g}\int_{\tau_0}^{\tau_1} \sqrt{x}\,\sqrt{x'^2 + y'^2 + z'^2}\, d\tau.$$

Euler's Equations take the form:

$$m\frac{d}{d\tau}\frac{\sqrt{x}\,x'}{\sqrt{S}} = \frac{\sqrt{S}}{\sqrt{x}}, \qquad \frac{d}{d\tau}\frac{\sqrt{x}\,y'}{\sqrt{S}} = 0, \qquad \frac{d}{d\tau}\frac{\sqrt{x}\,z'}{\sqrt{S}} = 0.$$

From the last two of these equations follows:

$$ay' + bz' = 0,$$
$$ay + bz + c = 0,$$

i.e. the path lies in a vertical plane. Let this be the (x, y)-plane.

The first equation now becomes, if we set the parameter $\tau = x$:

$$\frac{d}{dx}\frac{\sqrt{x}}{\sqrt{1+q}} = \frac{\sqrt{1+q}}{2\sqrt{x}}, \qquad q = \left(\frac{dy}{dx}\right)^2.$$

Hence

$$x\frac{dq}{dx} = -q(1+q)$$

$$\int\frac{dx}{x} = -\int\frac{dq}{q(1+q)}$$

$$\log x = \log\frac{1+q}{q} + \log a, \qquad \frac{x}{a} = \frac{1+q}{q};$$

$$\frac{dy}{dx} = \pm\frac{\sqrt{a}}{\sqrt{x-a}},$$

$$(y-b)^2 = 4a(x-a),$$

the equation of a parabola whose directrix is the axis of y.

The time is given by the equation:

$$t = \int_{\tau_0}^{\tau_1}\sqrt{\frac{\frac{1}{2}m(x'^2+y'^2+z'^2)}{mgx}}\,d\tau = \frac{1}{\sqrt{2g}}\int_{x_0}^{x_1}\sqrt{\frac{1+y'^2}{x}}\,dx$$

$$= \frac{1}{\sqrt{2g}}\int_{x_0}^{x_1}\frac{dx}{\sqrt{x-a}} = \sqrt{\frac{2}{g}}\left(\sqrt{x_1-a} - \sqrt{x_0-a}\right).$$

In the foregoing solution it has been tacitly assumed that x steadily increases from x_0 to x_1, since otherwise the choice of $\tau\,(=x)$ would not be permissible. If x steadily decreases, it is legitimate to set $-x=\tau$, the formulas now being correspondingly modified. But if x decreases for a time and then increases, it is still possible to transform to x as the variable of integration; only the interval $\tau_0 \leqq \tau \leqq \tau_1$ must be divided into two intervals, in one of which x is decreasing (here, $\tau = -x + \text{const.}$) and in the other of which x is increasing; but the derivative, $dx/d\tau$, will not be continuous in this case.

Finally, we have excluded the solution of Euler's Equations:

$$y' = 0, \qquad z' = 0.$$

It is this solution that corresponds to a vertical path: $y = y_0$, $z = z_0$.

Example 2. Consider the path of a particle on a smooth surface, $x = f(u, v)$, $y = \phi(u, v)$, $z = \psi(u, v)$, when no forces act except the reaction of the surface. Here the potential energy is constant, and we may set $U = 0$. Let $m = 2$, $h = 1$. Then

$$T = \frac{ds^2}{dt^2} = E\frac{du^2}{dt^2} + 2F\frac{du}{dt}\frac{dv}{dt} + G\frac{dv^2}{dt^2},$$

and Hamilton's Integral becomes:

$$\int_{t_0}^{t_1} (Eu^2 + 2F\dot{u}\dot{v} + Gv^2)\, dt.$$

Thus Hamilton's Principle leads to the equations of the geodesics in the form:

$$2\frac{d}{dt}(E\dot{u} + Fv) = E_u\dot{u}^2 + 2F_u uv + G_u\dot{v}^2,$$

$$2\frac{d}{dt}(Fu + G\dot{v}) = E_v u^2 + 2F_v uv + G_v\dot{v}^2.$$

If, on the other hand, we use the Principle of Least Action, we have:

$$\mathfrak{A} = \sqrt{2}\int_{\tau_0}^{\tau_1} \sqrt{S}\, d\tau,$$

where

$$S = Eu'^2 + 2Fu'v' + Gv'^2.$$

The condition that the action be stationary is here:

$$\frac{d}{d\tau}\frac{Eu' + Fv'}{\sqrt{S}} = \frac{E_u u'^2 + 2F_u u'v' + G_u v'^2}{2\sqrt{S}},$$

$$\frac{d}{d\tau}\frac{Fu' + Gv'}{\sqrt{S}} = \frac{E_v u'^2 + 2F_v u'v' + G_v v'^2}{2\sqrt{S}}.$$

The two forms come together on the basis of the relation which defines the time:

$$T\frac{dt^2}{d\tau^2} = S.$$

Here,

$$T = U + h = 1,$$

and so

$$dt = \sqrt{S}\, d\tau.$$

EXERCISES

1. *Carry through the details in Example 1 of the text when x sometimes decreases and sometimes increases. Show that the path* is still given by the same equation, but the formula for the time is modified. Check all results by comparison with the elementary treatment of the problem, *Introduction to the Calculus*, p. 384.

2. Carry through the solution of Example 1, setting $r = y$.

3. A particle moves under the action of a central force attracting according to the law of the inverse square of the distance. Prove the theorem of equal areas:

$$r^2 \frac{d\theta}{dt} = \text{const.}$$

CHAPTER XVIII

THERMODYNAMICS. ENTROPY

The object of this chapter * is to give the layman those physical pictures which enter in the conception of *Entropy*, and to show how this conception attains its simplest and most natural expression in the language of mathematics, — namely, as a line integral which is independent of the path of integration.†

The specialist in physics does well to trace step by step the physical phenomena, tested by laboratory experiment in the broadest sense of the term, which led up to the introduction of entropy as a line integral; but it is a mistake to assume that this inductive method affords the sole access to the conception, and it may even be questioned whether this approach is the best for the physicist. Why throw a smoke screen over the mountain he is to ascend? The climb is hard enough at best; why not let him have a good view of the glorious summit before he starts?

1. Reversible Changes and the (v, p)-Diagram. Imagine a hollow brass cylinder, 1 cm. in cross-section, closed at one end and provided with a piston. Let a quantity of air ‡ (1 gr., say) be present in the cylinder, and let the temperature, t, be the same throughout the system and hence, in particular, constant throughout the air. Let the pressure of the air per sq. cm. be denoted by p, and its total volume by v. Then p will be the total force exerted by the air on the piston, and the altitude of the part of the cylinder occupied by the air will be v.

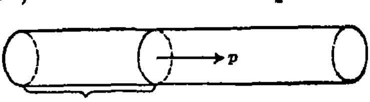


FIG. 111

* In the final editing of this chapter I am indebted to Mr. B. O. Koopman for a number of helpful suggestions.

† In this chapter we are taking the classical point of view, and not that of statistical thermodynamics — interesting and important though that may be.

‡ The whole discussion applies equally well to dry steam, the temperature being, of course, high enough to prevent condensation. More generally, it applies to any perfect gas.

Let the piston be moved, in or out, to a new position. Then the temperature of the air will change, and will not even be the same throughout at a given instant. If, however, the piston is moved slowly, and if the temperature of the walls of the chamber is nearly the same at all points at any given instant, the temperature of the air will also be nearly uniform at each instant, and we may contemplate the ideal case in which it is actually *uniform at each instant*, though different at different times.

Although this condition can never be accurately fulfilled in practice, nevertheless it is a sufficiently close approximation to the actual state of affairs in the most important cases, and physicists do not hesitate to begin the treatment of Thermodynamics by laying down the above hypothesis. We shall refer to it as the Fundamental Hypothesis and assume that it is fulfilled in all that follows.

Just as, in a problem of analytic mechanics, we begin by *isolating the system*, so here it is the air in the cylinder which is the material system under consideration. The brass of the cylinder is merely a means of keeping the air in place and transferring heat to or from it. For we can heat or cool the air, for any given value of v, *i.e.* for any given position of the piston, by allowing heat to flow in or flow out through the walls of the chamber.

We are now in a position to describe the physical picture which illustrates the thermodynamical phenomena in question. Let the piston be moved in or out. Mathematically this means that v varies, decreasing or increasing. Let heat be transferred to the air through the walls of the chamber, and let p denote the pressure corresponding to any given value of v. The pair of values (v, p) can be represented graphically by a point,* whose coordinates are v and p; and conversely, to an arbitrary pair of values of v and p, *i.e.* to any arbitrary point (v, p), corresponds a definite state of the body of air, in which the volume has the prescribed value v and the pressure is brought, by suitably applying heat, to the prescribed pressure, p.

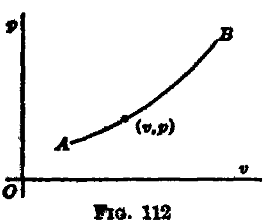


FIG. 112

Continuous Changes of State. We illustrate what is here meant by two simple examples and proceed then to the general case.

* The pair of values is read in physics in the order p, v, and physicists speak of the "(p, v)-diagram," although they plot the point as indicated. It would have been better to conform to this usage.

(*a*) *Isothermal Changes.* Suppose the brass cylinder containing the air to be immersed in a huge tank of water at constant temperature. Let the piston be slowly drawn out. Then the Fundamental Hypothesis will be nearly realized, and we are led to the conception of the ideal case, in which it is actually realized. The law connecting the pressure with the volume as the body of air undergoes this continuous change of state is:

$$pv = C \qquad \text{or} \qquad pv = p_0 v_0,$$

and is known as *Boyle's Law.**

Thus a curve (here, a hyperbola) is described in the (v, p)-plane. The process is accompanied by a steady influx of heat into the air from the walls of the cylinder. If, now, the piston is gradually pushed back, the point (v, p) will describe the same curve in the opposite sense, and a quantity of heat will be given out just equal to the amount taken up in the direct process. The process is, therefore, described as *reversible.*

(*b*) *Adiabatic Changes.* Suppose the brass cylinder and piston (now thought of as thin and thus having small volume compared with the body of air) to be insulated in asbestos, so that practically no heat enters or escapes in a considerable interval of time, as the piston is gradually drawn out. Thus we are led to the conception of a chamber whose walls are absolutely *adiabatic, i.e.* impervious to any transfer of heat. The change of state that now arises under the assumption of the Fundamental Hypothesis is known as an *adiabatic change*, and the law connecting p and v is, for air, approximately †

$$pv^{1.4} = C \qquad \text{or} \qquad pv^{1.4} = p_0 v_0^{1.4}.$$

The process is *reversible*; *cf.* Fig. 113, p. 459.

An adiabatic change can be realized approximately in a bicycle pump when the piston is rapidly pushed down, the pump and air being initially at a uniform temperature.

(*c*) *The General Case.* Consider now an arbitrary curve in the (v, p)-plane. We may imagine that the piston is slowly moved and

* We may at this point recall the *Law of Charles*, which asserts that

$$\frac{p}{p_0} = \frac{T}{T_0},$$

where v is constant and T is the absolute temperature, § 2.

† The truth of this statement will be proved later on the basis of more fundamental physical facts.

at the same time heat is suitably poured into the body of air, or extracted from it, so that the pressure, p, and the volume, v, are always connected by the law indicated by the curve, the Fundamental Hypothesis being assumed to hold throughout. Thus at any point of the curve the quantity of heat, measured in calories, and taken algebraically (positive or negative), which has been required since the beginning of the process to maintain the pressure as prescribed by the curve, has a definite value, Q. If the process be reversed, the curve being now described in the opposite sense, the amount of heat, Q', required will be the negative of Q:

$$Q' = -Q.$$

2. The First Law of Thermodynamics. *Work.* The work, W, done by the body of air on the piston is given by the equation (*cf.* Chap. XI, §§ 1, 2)

$$W = \int_C p\, dv,$$

where C refers to the path in the (v, p)-plane.

Energy. Let the internal energy of the air be denoted by U. It is proportional to the absolute temperature,

$$T = t + 273,$$

where t is the temperature measured in degrees centigrade; *cf.* § 5, (11).

Finally, let the total amount of heat* in the body of air, measured in calories, be denoted by Q.

We are now in a position to state the First Law of Thermodynamics. It is this. Let the body of air experience an arbitrary continuous change of state, represented by a curve C in the (v, p)-plane. Then the heat that has flowed in, $Q - Q_0$, is proportional to the increment in the internal energy, $U - U_0$, plus the work, W, done by the gas on the piston, or

$$(1) \qquad Q - Q_0 = A\Big((U - U_0) + W\Big),$$

where A is a constant which will be discussed presently.

* In all that follows we are dealing, not with the total amount of heat in the body, but with an amount introduced into the body (or extracted from it) during a given process. It is convenient to think of this quantity as a *difference*, and a simple way to do this is to start with the idea of the total amount of heat in the body at any time.

Joule's Equivalent. The number of units of work required to raise the temperature of the unit mass of water one degree is known as *Joule's equivalent of heat* and is represented by J. In the c.g.s. system J has the value 427; *i.e.* the work done by gravity on one gramme of matter, when the latter is lowered 427 metres, is just sufficient to raise the temperature of one gramme of water one degree centigrade.*

The constant, A, is the reciprocal of J,

$$A = \frac{1}{J}.$$

The Law admits extensions of the most varied character. Our object here, in confining ourselves to this simple case of the air in the cylinder and of reversible processes, is to give an absolutely concrete picture of what goes on in a typical case, and to set forth the mathematical methods which apply, not merely here, but in the more complex cases.

3. Differentials. The First Law of Thermodynamics is often expressed by physicists in the form:

$$(2) \qquad dQ = A(dU + dW) \qquad \text{or} \qquad = A(dU + p\,dv).$$

What do these differentials mean? Are they the differentials of functions, and if so, what are the independent variables? Or are they merely approximations for increments, — the true equation being

$$(3) \qquad \Delta Q = A(\Delta U + \Delta W)$$

and if so, how are these increments taken?

The crux of the whole matter, in either case, is this question of what the *independent variables* are. The answer is that *sometimes these are two in number, and sometimes only one.* Thus in the case of the energy, U, this quantity is completely determined by v and p. If the gas, starting from an initial state (v_0, p_0), passes continuously through any changes of state represented by a curve C in the (v, p)-plane, which comes back to (v_0, p_0) and thus closes, the energy at the end of the trip will be exactly the same as the energy at the beginning. Hence U is a function of the two independent variables, v and p, and its differential, dU, has the value

* For example, if a rain drop were to fall *in vacuo* from a distance of 427 metres (or nearly a quarter of a mile) into a pail of water, the heat generated by the shock would be just sufficient to raise the temperature of the rain drop one degree centigrade.

(4) $$dU = \frac{\partial U}{\partial v} dv + \frac{\partial U}{\partial p} dp,$$

where $dv = \Delta v$ and $dp = \Delta p$ are the arbitrary independent increments of these variables.

On the other hand, the work, W, can never be a function of v and p, regarded as independent variables. It depends essentially on the choice of the curve, C. What does dW mean in this case? Clearly, the curve, C, must come first. Then we have a one-dimensional range of values, and the length of the arc, s, of C affords a natural choice of the independent variable. Or, if C is nowhere vertical, we may take v as the independent variable. In either case we have

$$dW = p\,dv.$$

In the latter case, $dv = \Delta v$; in the former, $dv = D_s v \Delta s$.

Precisely the same remarks apply to Q. Like W, Q depends on the choice of C and, after this choice has been made, becomes then a function of a single variable, as s or v.

We are now in a position to answer the question of the meaning of the differentials in equation (2). First, an arbitrary point, (v, p), is chosen; secondly, an arbitrary curve, C, is passed through this point; thirdly, the quantities, U, Q, and W are considered along this curve, under the assumption that the air experiences a continuous change of state, as defined by the Fundamental Hypothesis and represented by C. Thus U becomes a function of the single variable s, — the length of the arc of C, — and the same is true of v and p. Equation (2) is, however, still true, for it is a fundamental property of the total differential of a function that equation (B) of Chap. V, § 5 is true, *no matter what the independent variables may be.* Thus equation (2) really means no more and no less than that

(5) $$D_s Q = A(D_s U + D_s W),$$

where the derivatives are taken along C. That, however, these derivatives should be connected by any such relation, and that the same equation (2) or (5) should hold for *all* curves C through a given point, is in no wise evident, either mathematically or physically. *It is in this equation that the physical law finds its complete expression.*

What of the other view, that (2) is a near-equation for (3)? Here, the differentials seem self-explanatory: they are close approximations for the increments, and why worry?

How superficial this view is, becomes evident when we ask what the increments are. How does the gas pass from the state represented by the point (v, p) to the state represented by the point $(v + \Delta v, p + \Delta p)$? For there is an infinite number of possible paths connecting these points, and *some* path must be followed, or we have no definite physical picture before us. But as soon as we introduce a path, — a curve C through (v, p), — we have all the preliminary physical pictures of the earlier explanation. And this is, in fact, the answer. Equation (3) requires precisely the same physical setting described in the first explanation of equation (2). When all this has been done, we can then divide (3) through by Δs, let Δs approach 0, and thus deduce equation (5). Not even yet have we arrived at (2); for it is (5), a relation between *derivatives*, which expresses the physical facts directly, and the transition to (2) is purely mathematical.

There is no short cut, no self-explanatory method, whereby (2) is written down without the intervention of any curve C and the derivatives and differentials pertaining to it, as Pallas Athene sprang, panoplied, from the head of Zeus.

4. In Particular, the Differential dQ. From equations (2) and (4) it follows that

(6) $$dQ = X\,dv + Y\,dp,$$

where $$X = A\frac{\partial U}{\partial v} + Ap, \qquad Y = A\frac{\partial U}{\partial p}.$$

Here, X and Y are functions of the two independent variables, v and p, but the expression

$$X\,dv + Y\,dp$$

is not an exact differential. Physically, this is clear from the fact that, when the integral

$$\int X\,dv + Y\,dp$$

is extended over a closed path, it represents the heat that has been transferred to the gas in the process, and this is not in general zero, for the work is not in general zero. Mathematically,

$$\frac{\partial X}{\partial p} - \frac{\partial Y}{\partial v} = A,$$

and the right-hand side of this equation would have to be zero, if (6) were an exact differential; Chap. XI, § 7.

May it not be possible, however, to multiply equation (6) through by such a function of v and p that the new right-hand side does become an exact differential? The answer to this question is affirmative. It comprises a physical law of the greatest importance and it leads to the introduction of a new physical quantity, — *entropy*. It is as follows*: *The absolute temperature, T, a function of the two independent variables v and p, yields through its reciprocal an integrating factor; i.e.*

$$\frac{X\,dv + Y\,dp}{T}$$

is an exact differential.

Thus it appears that the integral

$$(7)\qquad \int_{(v_0,\,p_0)}^{(v,\,p)} \frac{X\,dv + Y\,dp}{T},$$

taken along any path connecting the fixed point (v_0, p_0) with the arbitrary point (v, p), is independent of the particular path chosen, and consequently defines a function of the two independent variables, v and p. It is reasonable to expect that this function,

$$(8)\qquad \mu = \int_{(v_0,\,p_0)}^{(v,\,p)} \frac{dQ}{T} = \int_{(v_0,\,p_0)}^{(v,\,p)} \frac{X\,dv + Y\,dp}{T},$$

should have a physical meaning, and this is, in fact, the case. The quantity μ (or μ + a constant) is known as the *entropy* of the body.

By the first of the integrals (8) is meant nothing more or less than

$$\int_0^l \frac{D_s Q}{T}\,ds, \qquad \text{or} \qquad \int_C \frac{D_s Q}{T}\,ds,$$

this integral being extended along the curve C of the (v, p)-plane; Chap. XI, § 3.

* We are, of necessity, omitting the physical considerations which lead inductively to this statement, namely, the discussion of the efficiency of reversible and non-reversible "heat engines," the definition of the thermodynamic scale of temperature, etc. For, as was said at the beginning of the chapter, our object is rather to show how the final form of the law receives mathematical expression by means of the idea of line integrals and exact differentials. The reader who wishes to inform himself concerning the physical phenomena which lead up to the law may refer to any of the standard works, *e.g.* Buckingham, *Theory of Thermodynamics*; Blondlot, *Introduction à l'étude de la thermodynamique*; Poincaré, *Thermodynamique.*

We have here a notable example of the interplay between mathematics and physics. Expressions of the form (6), so-called "inexact differentials," were studied in an earlier chapter from a purely mathematical standpoint, and it was shown analytically that an "integrating factor" always exists; Chap. XIV, § 20. Here, the function (8) thus obtained has a physical meaning of prime importance.

5. The Entropy of a Perfect Gas. In the case of a perfect gas, as air or dry steam, we have, on combining the laws of Boyle and Charles,

$$(9) \qquad pv = \alpha p_0 v_0 T, \qquad \alpha = \tfrac{1}{273},$$

where $T = 273 + t$ denotes the absolute temperature, t being the temperature centigrade, and v_0 and p_0 are values of v and p corresponding to a temperature of 0° centigrade.

Thus any two of the three quantities p, v, T may be taken as the independent variables, and the third is determined as a function of these by (9). Beside the (v, p)-diagram we have now, by a purely mathematical transformation, a (t, v)-diagram and a (t, p)-diagram.

In the case of the (t, v)-diagram the First Law of Thermodynamics, § 3, (2) here assumes the form

$$(10) \qquad dQ = c\,dT + l\,dv,$$

$$c = A\frac{\partial U_{vT}}{\partial T}, \qquad l = A\frac{\partial U_{vT}}{\partial v} + Ap.$$

The coefficient c is constant (or nearly so); it is the *specific heat for constant volume*, or the number of calories required to raise one gramme of the gas one degree.* Hence

$$(11) \qquad U = \frac{c}{A}T.$$

In the case of the (t, p)-diagram, the First Law of Thermodynamics assumes the form

$$(12) \qquad dQ = C\,dT + h\,dp,$$

$$C = A\left(\frac{\partial U}{\partial T} + p\frac{\partial v}{\partial T}\right), \qquad h = Ap\frac{\partial v}{\partial p}.$$

From (9) it follows, since the independent variables are here T and p, that

$$p\frac{\partial v}{\partial T} = \alpha p_0 v_0.$$

* Moreover, U does not change with v when T is constant, and so $\partial U_{vT}/\partial v = 0$. Thus $l = Ap$. l is the *latent heat* of expansion.

Hence

(13) $$C - c = \alpha A p_0 v_0.$$

Computation of dQ/T. For a perfect gas, equation (6) of § 4 yields the following. Since

$$T = \frac{A}{C-c} pv, \qquad U = \frac{c}{C-c} pv,$$

we have:

$$X = A\left(\frac{\partial U}{\partial v} + p\right) = A\left(\frac{c}{C-c} p + p\right) = \frac{AC}{C-c} p;$$

$$Y = A\frac{\partial U}{\partial p} = \frac{Ac}{C-c} v.$$

Hence

(14) $$\frac{dQ}{T} = C\frac{dv}{v} + c\frac{dp}{p}.$$

The entropy is given by the integral:

$$\mu - \mu_0 = \int_{(v_0, p_0)}^{(v,p)} \frac{dQ}{T} = \int_{(v_0, p_0)}^{(v,p)} C\frac{dv}{v} + c\frac{dp}{p},$$

or

(15) $$\mu - \mu_0 = \log(v^C p^c) \Big|_{(v_0, p_0)}^{(v,p)},$$

(16) $$\mu = \log(v^C p^c) + \text{const.}$$

The Curves $T = \text{const.}$, $\mu = \text{const.}$ From (9) it follows that the curves $T = \text{const.}$ are the hyperbolas:

(17) $$pv = \text{const.}$$

For air, the ratio C/c has the value 1.4. Hence the adiabatics, or curves of constant entropy, are the family

(18) $$pv^{1.4} = \text{const.}$$

Thus it appears that the latter curves are steeper than the former.

Two Deductions. (*i*) If we take as the curve in the (v, p)-plane an *adiabatic*, *i.e.* a curve corresponding to a change in which no heat is absorbed or given out, then $\frac{dQ}{ds} = 0$ at every point of the curve, and hence

(19) $$\mu - \mu_0 = \int_{(v_0, p_0)}^{(v,p)} \frac{dQ}{T} = 0.$$

Consequently, *the curves of constant entropy are adiabatics.*

(*ii*) If we take as the curve in the (v, p)-plane an *isothermal, i.e.* a curve for which $T = T_1 = \text{const.}$, then

$$(20)\qquad \mu - \mu_0 = \int_{(v_0, p_0)}^{(v, p)} \frac{dQ}{T_1} = \frac{Q - Q_0}{T_1}.$$

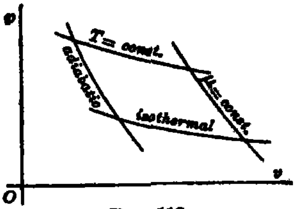


Fig. 113

Physical Definition of Entropy. From (19) and (20) it follows that it is possible to define the *entropy* as a function of (v, p) such that *it is constant along an adiabatic, and its change along an isothermal,* $T = T_1$, *is equal to*

$$\frac{Q - Q_0}{T_1}.$$

CHAPTER XIX

DEFINITE INTEGRALS AND THE GAMMA FUNCTION

1. The Definite Integral as a Function of a Parameter. Leibniz's Rule. Consider the definite integral

$$\int_a^b f(x, \alpha)\, dx,$$

where $f(x, \alpha)$, for a fixed value of α, is a continuous function of x in the interval $a \leqq x \leqq b$. Since the integral has a definite value for each value of α, it is a function of α:

(1) $$\phi(\alpha) = \int_a^b f(x, \alpha)\, dx.$$

We will require, furthermore, that $f(x, \alpha)$ be a *continuous* function of the *two independent variables* (x, α) throughout the region

(2) $$a \leqq x \leqq b, \qquad A \leqq \alpha \leqq B.$$

It follows that $\phi(\alpha)$ is a continuous function of α. The proof is immediate when we visualize the geometric picture. The surface

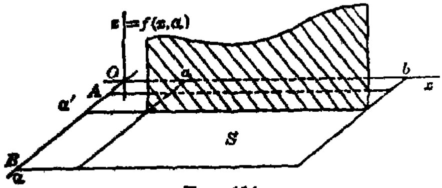


Fig. 114

$$z = f(x, \alpha)$$

is cut by the plane $\alpha = \alpha'$ in a curve,

$$z = f(x, \alpha'), \qquad \alpha = \alpha',$$

and the area under this curve represents the value of the definite integral formed for this value of α:

$$\phi(\alpha') = \int_a^b f(x, \alpha')\, dx.$$

If α' is changed by a small amount, the plane shifts slightly, and the area also changes but slightly. This is precisely the condition that $\phi(\alpha)$ be continuous.

Differentiation of $\phi(\alpha)$. Apply the definition of a derivative. Since $\Delta\alpha$ is a constant with respect to the variable of integration, we can divide by it under the sign of integration, and thus we have:

$$\frac{\phi(\alpha_0 + \Delta\alpha) - \phi(\alpha_0)}{\Delta\alpha} = \int_a^b \frac{f(x, \alpha_0 + \Delta\alpha) - f(x, \alpha_0)}{\Delta\alpha}\, dx. \tag{3}$$

As $\Delta\alpha$ approaches 0, the integrand approaches the limit $\partial f/\partial\alpha = f_\alpha(x, \alpha_0)$, and thus it would appear that the limit of the left-hand side, or $D_\alpha\phi(\alpha)$, has the value

$$\frac{d\phi}{d\alpha} = \int_a^b \frac{\partial f}{\partial\alpha}\, dx. \tag{4}$$

This is, in fact, the case when we impose the further condition that the derivative $\partial f/\partial\alpha$ exist, and that the function $f_\alpha(x, \alpha)$ be continuous in the two-dimensional region (2). But the proof is not so simple as one might think, for what we want is the limit approached by the integral of $\Delta_\alpha f/\Delta\alpha$, and there is no reason to suppose that this is the same as the integral of the limit approached by $\Delta_\alpha f/\Delta\alpha$. In other words, it is a question of reversing the order of the operations in a double limit — the process of differentiation and that of integration.

It is not difficult, however, to avoid the fallacy just pointed out. By the Law of the Mean,

$$f(x, \alpha_0 + \Delta\alpha) - f(x, \alpha_0) = \Delta\alpha f_\alpha(x, \alpha_0 + \theta\Delta\alpha), \qquad 0 < \theta < 1. \tag{5}$$

Hence the right-hand side of (3) becomes

$$\int_a^b f_\alpha(x, \alpha_0 + \theta\Delta\alpha)\, dx. \tag{6}$$

We can form readily a geometric picture of the value of this integral. θ is, of course, a function of all the variables in sight, namely, $x, \alpha_0, \Delta\alpha$. The equations

$$z = f_\alpha(x, \alpha), \qquad \alpha = \alpha_0 + \theta\Delta\alpha$$

represent a space curve lying on the surface $z = f_\alpha(x, \alpha)$ and comprised between the planes $\alpha = \alpha_0$ and $\alpha = \alpha_0 + \Delta\alpha$. The projection of this space curve on the plane $\alpha = \alpha_0$ is a continuous curve,

$$z = f_\alpha(x, \alpha_0 + \theta\Delta\alpha),$$

since the left-hand side of (5) is continuous, and the area under this curve, or the integral (6), is thus seen to differ but slightly from the value of the integral (4) when $\Delta\alpha$ is numerically small. This completes the proof.* Equation (4) is known as *Leibniz's Rule.*

Variable Limits of Integration. The limits of integration, a and b, may depend on α, or more generally they may vary in any manner. Let S be a region of the (x, α)-plane as indicated in the figure. Let α have any value α' in the interval (A, B), and let the segment

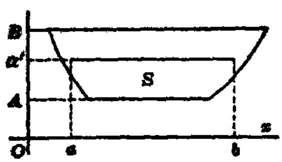


FIG. 115

$$a \leqq x \leqq b, \qquad \alpha = \alpha'$$

lie in S. Then the integral (1) is a function of the three independent variables (α', a, b). Dropping the prime we write

$$(7) \qquad \phi(\alpha, a, b) = \int_a^b f(x, \alpha)\, dx.$$

The function $f(x, \alpha)$, together with its partial derivative $f_\alpha(x, \alpha)$, shall be continuous throughout S. It follows that $\phi(\alpha, a, b)$ is a continuous function of all three arguments, as is seen from the geometric representation corresponding to Fig. 114.

The function $\phi(\alpha, a, b)$ admits partial derivatives whose values are:

$$(8) \quad \frac{\partial\phi}{\partial\alpha} = \int_a^b f_\alpha(x, \alpha)\, dx, \qquad \frac{\partial\phi}{\partial a} = -f(a, \alpha), \qquad \frac{\partial\phi}{\partial b} = f(b, \alpha).$$

In particular, a and b may be made to depend on α:

$$(9) \qquad a = \psi(\alpha), \qquad b = \omega(\alpha),$$

where ψ and ω are continuous together with their first derivatives, and the curves (9) lie in S. We then have a generalization of Leibniz's Theorem which is embodied in the formula:

$$(10) \quad \frac{d}{d\alpha}\int_a^b f(x, \alpha)\, dx = \int_a^b \frac{\partial f}{\partial \alpha}\, dx + f(b, \alpha)\frac{db}{d\alpha} - f(a, \alpha)\frac{da}{d\alpha}.$$

* The arithmetic form of the proof depends on the *uniform continuity* of the function $f_\alpha(x, \alpha)$; *cf.* Goursat-Hedrick, *Mathematical Analysis*, vol. I, § 97, or the author's *Funktionentheorie*, vol. I, Chap. III, § 8.

EXERCISES

Verify Leibniz's Rule by evaluating explicitly the integrals in each of the following cases:

1. $\int_0^1 e^{-\alpha x}\,dx.$ 2. $\int_{-\alpha}^{\alpha}\frac{dx}{1+\alpha^2+x}.$ 3. $\int_0^{\alpha}\sqrt{\alpha^2-x^2}\,dx.$

2. Several Parameters and Multiple Integrals. The foregoing theorems admit generalization to the case that the integrand depends on several parameters,

$$\int_a^b f(x,\,\alpha,\,\beta,\,\gamma,\,\cdots)\,dx$$

and also to that of multiple integrals, *e.g.*

$$\iiint\limits_V f(x,\,y,\,z\,;\;\alpha,\,\beta,\,\gamma,\,\cdots)\,ax\,dy\,dz.$$

If the integrand is continuous in all the arguments, the function defined or represented by the integral will be continuous in all the parameters; and if, in addition, the partial derivative $\partial f/\partial\alpha$ is continuous in all the arguments, the function will admit a partial derivative given by the earlier formula:

$$\text{(11)}\quad \frac{\partial}{\partial\alpha}\int_a^b f(x,\,\alpha,\,\beta,\,\gamma,\,\cdots)\,dx=\int_a^b\frac{\partial f}{\partial\alpha}\,dx+f(b,\,\alpha,\,\beta,\,\gamma,\,\cdots)\frac{\partial b}{\partial\alpha}$$
$$-f(a,\,\alpha,\,\beta,\,\gamma\,\cdots)\frac{\partial a}{\partial\alpha};$$

$$\text{(12)}\quad \frac{\partial}{\partial\alpha}\iiint\limits_V f\,dV=\iiint\limits_V\frac{\partial f}{\partial\alpha}\,dV.$$

In the case of multiple integrals we assume that the region of integration is fixed. Cases arise in hydromechanics, in which the region varies with the parameters, but the treatment does not belong to the elements of the Calculus.

Example. The *potential*, u, of a continuous volume distribution of matter is defined by the formula, Chap. IV, § 5:

$$u=\iiint\limits_V\frac{\rho\,dV}{r},$$

where the density, ρ, is a continuous function of the coordinates (a, b, c) of a variable point Q in V. The point P at which the potential is measured has the coordinates (x, y, z); it is exterior to V, and

$$r^2 = (x-a)^2 + (y-b)^2 + (z-c)^2.$$

The force with which this distribution attracts a unit particle at P has for its components along the coordinate axes the values

$$X = \frac{\partial u}{\partial x}, \qquad Y = \frac{\partial u}{\partial y}, \qquad Z = \frac{\partial u}{\partial z}.$$

These derivatives are given by the integrals

$$\frac{\partial u}{\partial x} = -\iiint_V \frac{\rho(x-a)}{r^3}\,dV, \qquad \text{etc.}$$

EXERCISE

The potential of a surface distribution is given by the surface integral

$$u = \iint_S \frac{\sigma\,dS}{r},$$

where σ denotes the density. Find the components of the attraction at a point not lying on the surface.

3. Improper Integrals. If we evaluate the following integrals by the usual rule, we find:

$$(1) \qquad \int_0^1 \frac{dx}{\sqrt{x}} = 2\sqrt{x}\Big|_0^1 = 2; \qquad \int_{-1}^{1} \frac{dx}{x^2} = -\frac{1}{x}\Big|_{-1}^{1} = -2.$$

The first result looks reasonable, for the area under the curve $y = 1/\sqrt{x}$ in the interval $(0, 1)$ might well be 2 units; but the second result is absurd, for the curve lies above the axis of x throughout the whole interval, and so the area under the curve cannot be negative.

Clearly, then, such integrals — integrals whose integrands do not remain finite — cannot be treated by the same rules as ordinary definite integrals. They are examples of *improper integrals* (*uneigentliche Integrale*), and under this class is to be included one further case, in which the integration is extended over an infinite interval. We will, in fact, begin with this latter case.

The Improper Integral $\int_c^\infty f(x)\,dx$. Let $f(x)$ be continuous in the interval $c \leqq x < \infty$. Form the integral,

$$\int_c^x f(x)\,dx. \tag{2}$$

Let x increase without limit. If this integral approaches a limit, we say that

$$\int_c^\infty f(x)\,dx \tag{3}$$

converges and assign to it as its *value* the limit approached:

$$\int_c^\infty f(x)\,dx = \lim_{x=\infty} \int_c^x f(x)\,dx.$$

If, however, the integral (2) approaches no limit, we say that the integral (3) *diverges*. No value is assigned to a divergent integral.*

Example.

$$\int_0^\infty e^{-x}\,dx = 1.$$

For,

$$\int_0^x e^{-x}\,dx = 1 - e^{-x} \qquad \text{and} \qquad \lim_{x=\infty} \int_0^x e^{-x}\,dx = 1.$$

Geometrically this means that the area under the curve

$$y = f(x) = e^{-x}$$

in the interval $(c, x) = (0, x)$ approaches a limit, namely 1, as x becomes infinite.

Functions defined by Improper Integrals. Consider the integral

$$\int_0^\infty \frac{\alpha\,dx}{\alpha^2 + x^2}. \tag{4}$$

* The student will perceive the close analogy between such improper integrals and infinite series. The integral (2) corresponds precisely to s_n (*Introduction to the Calculus*, Chap. XIV, § 2), and the integral (3) to the infinite series, $u_1 + u_2 + \cdots$.

If $\alpha \neq 0$, we have

(5) $$\int_0^x \frac{\alpha\, dx}{\alpha^2 + x^2} = \tan^{-1}\frac{x}{\alpha}.$$

Hence the integral (4) converges, and its value is $\pi/2$ when $\alpha > 0$, and $-\pi/2$ when $\alpha < 0$.

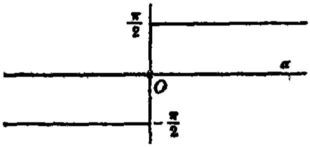


FIG. 116

The integral (4) also converges when $\alpha = 0$. Equation (5) no longer has a meaning; but the integral on the left of equation (5) has,* and its value is 0 for all values of x. Hence its limit is 0.

To sum up: The integral (4) converges for *all* values of α, and it defines a discontinuous function $\phi(\alpha)$:

(6) $$\phi(\alpha) = \int_0^\infty \frac{\alpha\, dx}{\alpha^2 + x^2} = \begin{cases} \frac{\pi}{2}, & 0 < \alpha; \\ 0, & \alpha = 0; \\ -\frac{\pi}{2}, & \alpha < 0. \end{cases}$$

It should be pointed out that no question of indeterminateness can arise concerning equation (5), since α is chosen first ($\alpha \neq 0$) and then held fast; *i.e.* α is a constant, and x is the variable.

4. Tests for Convergence. TEST BY DIRECT COMPARISON. INTEGRAND POSITIVE. *Let $\phi(x)$ be continuous and positive (or zero) in the interval $g \leqq x < \infty$, where g is any convenient fixed number in the interval* $c \leqq x < \infty$; *i.e.*, $g \geqq c$. *Let*

(1) $$\int_g^\infty \phi(x)\, dx$$

converge. If $f(x)$ is positive (or zero) and not greater than $\phi(x)$ at any point of the interval $g \leqq x < \infty$:

* The ordinary definition of a definite integral as given in the *Introduction*, Chap. XII, or in the present volume, Chap. XII, admits immediate extension to the case that the integrand is undefined in a finite number of points of the interval, provided it is continuous in all other points, and remains finite in the whole interval, *i.e.*

$$|f(x)| < G, \qquad a \leqq x \leqq b, \qquad G, \text{ a positive constant.}$$

$$0 \leqq f(x) \leqq \phi(x),$$

then the integral (3) *of* § 3 *converges.*

For,
$$\int_g^x f(x)\,dx \leqq \int_g^x \phi(x)\,dx \leqq \int_g^\infty \phi(x)\,dx = A.$$

Thus the integral (2), § 3, being equal to

$$\int_c^g f(x)\,dx + \int_g^x f(x)\,dx, \qquad g < x,$$

is seen to increase (or remain constant) as x increases, but never to exceed the constant value $A + B$, where B denotes the value of the first of these last two integrals. Hence, by the Fundamental Principle, *Introduction to the Calculus*, p. 391, this variable approaches a limit, and the theorem is proved.

The test is analogous to the Direct Comparison Test for Convergence in the case of infinite series whose terms are positive. Furthermore, as there, so here, a like test for divergence exists:

TEST FOR DIVERGENCE. *Let* $\phi(x)$ *be continuous in the interval* $g \leqq x < \infty$, *where* $g \geqq c$, *and let*

(2)
$$\int_g^\infty \phi(x)\,dx$$

diverge. If

$$f(x) \geqq \phi(x) \geqq 0, \qquad g \leqq x < \infty,$$

then the integral (3), § 3, *diverges.*

The proof is left to the student.

EXERCISES

1. Show by direct evaluation that

$$\int_g^\infty \frac{C\,dx}{x^p}, \qquad 0 < C, \qquad 0 < g, \qquad 1 < p,$$

converges.

2. Show in like manner that

$$\int_g^\infty \frac{C\,dx}{x}, \qquad 0 < C, \qquad 0 < g,$$

diverges.

Test for convergence or divergence each of the following integrals, using the theorems of the text and taking, as the comparison integral, (1) or (2), the integral of Question 1 or 2.

3. $\int_0^\infty \frac{dx}{\sqrt{1+x^4}}$. 4. $\int_1^\infty \frac{x\,dx}{\sqrt{x^4-1}}$. 5. $\int_0^\infty \frac{dx}{a^7+x^7}$.

6. $\int_1^\infty \frac{x\,dx}{\sqrt{1+x^5}}$. 7. $\int_1^\infty \frac{dx}{\sqrt{x^5-1}}$. 8. $\int_1^\infty \frac{x^6\,dx}{x^7-1}$.

5. Absolutely Convergent Integrals. If $f(x)$ is continuous in the interval $c \leqq x < \infty$, and if the integral

(1) $$\int_c^\infty |f(x)|\,dx$$

converges, then

(2) $$\int_c^\infty f(x)\,dx$$

converges. For, since

$$f(x) = |f(x)| - \{|f(x)| - f(x)\},$$

it follows that

$$\int_c^x f(x)\,dx = \int_c^x |f(x)|\,dx - \int_c^x \{|f(x)| - f(x)\}\,dx.$$

The first term on the right approaches a limit, by hypothesis. Now,

$$0 \leqq |f(x)| - f(x) \leqq 2|f(x)|.$$

Hence the second term also approaches a limit when x becomes infinite, and the theorem is proved.

When the integral (1) converges, the integral (2) is said to *converge absolutely*.

Example. $$\int_0^\infty \frac{\sin x\,dx}{1+x^2}.$$

Since $$\left|\frac{\sin x}{1+x^2}\right| \leqq \frac{1}{1+x^2}$$

and since the integral of this last function converges (either by the test of § 4 or by direct evaluation), the given integral is seen to converge.

6. The Limit Tests. We will still assume $f(x)$ to be continuous in the interval $c \leqq x < \infty$.

Test for Convergence. *If there exists a constant k greater than 1, such that $x^k f(x)$ approaches a limit:*

$$\lim_{x=\infty} x^k f(x) \quad \textit{exists}, \qquad 1 < k,$$

then the integral $\int\limits_c^\infty f(x)\,dx$ *converges.*

Let the above limit be denoted by L. Then it is clear that $x^k|f(x)|$ approaches a limit, namely $|L|$:

$$\lim_{x=\infty} x^k|f(x)| = |L|.$$

Let C be a positive constant greater than $|L|$. It follows that, for all values of x sufficiently large, this latter variable will be less than C, or

$$x^k|f(x)| < C, \qquad g \leqq x < \infty,$$

where g is a positive constant not less than c. Hence

$$|f(x)| < \frac{C}{x^k}, \qquad g \leqq x < \infty.$$

On setting $\phi(x) = C/x^k$ and applying to the integral

$$\int\limits_c^\infty |f(x)|\,dx$$

the test of § 4, this integral is seen to converge Hence the given integral converges absolutely.

A corresponding test for divergence exists, but it must be stated and applied with care.

Test for Divergence. *If*

(*i*) $\lim\limits_{x=\infty} xf(x)$ *exists*, — (denote it by L);

(*ii*) *and if this limit is not zero,* $L \neq 0$;

then the integral $\int\limits_c^\infty f(x)\,dx$ *diverges.*

Also, if $xf(x)$ becomes infinite, i.e. if

$$\lim_{x=\infty} xf(x) = +\infty \quad \textit{or} \quad = -\infty,$$

the above integral diverges.

Suppose that $L > 0$. Let a positive constant C be chosen less than L:

$$0 < C < L.$$

Then the variable $xf(x)$ will become and remain greater than C:

$$xf(x) > C, \qquad g \leqq x < \infty, \quad (g > 0).$$

Hence

$$f(x) > \frac{C}{x},$$

and the integral is seen to diverge by the test of § 4.

If $L < 0$, let $f(x) = -F(x)$. The proof of divergence can now be given in a similar manner.

When $xf(x)$ becomes positively infinite, any positive constant C can be chosen and the proof given as above. Finally, the case that $xf(x)$ becomes negatively infinite can be referred to the case just considered by setting $f(x) = -F(x)$.

Remark. If $L = 0$, we cannot infer either convergence or divergence, as the following examples show. In each case, $L = 0$.

$$(a) \quad \int\limits_2^\infty \frac{dx}{x \log x}. \qquad \text{Here} \quad \int\limits_2^x \frac{dx}{x \log x} = \log\log x - \log\log 2,$$

and thus this integral diverges.

$$(b) \quad \int\limits_2^\infty \frac{dx}{x(\log x)^2}. \qquad \text{Here} \quad \int\limits_2^x \frac{dx}{x(\log x)^2} = -\frac{1}{\log x}\Bigg|_2^x,$$

and this integral converges.

The tests of this paragraph are analogous to the Test-Ratio Test for the convergence of series, *Introduction to the Calculus*, p. 394.

Finally, the case of the improper integral

$$\int\limits_{-\infty}^{c} f(x)\,dx,$$

where $f(x)$ is continuous in the interval $-\infty < x \leqq c$, can be treated in a similar manner, or referred back to the above integral by means of the change of variable, $x = -t$.

Example 1. $\int\limits_0^\infty e^{-x^2}\,dx.$ Here, any value of $k > 1$ leads to the proof of convergence. Set, for simplicity, $k = 2$:

$$x^k f(x) = x^2 e^{-x^2}.$$

By the method of Chap. X, § 3, it is shown that this function approaches a limit, and hence the integral converges.

Example 2. $\displaystyle\int_0^\infty \frac{\cos x\,dx}{\sqrt{1+x^3}}.$ Here, any value of k between 1 and $\frac{3}{2}$ can be used, for

$$x^k \frac{\cos x}{\sqrt{1+x^3}} = \frac{\cos x}{x^{3/2-k}\sqrt{1+1/x^3}}.$$

Thus the integral converges.

EXERCISES

1. Show that the integral $\displaystyle\int_1^\infty \frac{dx}{\sqrt{x^4-1}}$ converges.

2. Show that the integral $\displaystyle\int_0^\infty \frac{x\,dx}{\sqrt{1+x+x^4}}$ diverges.

Test the following integrals for convergence or divergence, using each time the *simplest* method.

3. $\displaystyle\int_0^\infty \frac{dx}{\sqrt{1+x^3}}.$ 4. $\displaystyle\int_0^\infty \frac{x\,dx}{\sqrt{1+x^3}}.$ 5. $\displaystyle\int_0^\infty \frac{x\,dx}{\sqrt{1+x^5}}.$

6. $\displaystyle\int_0^\infty e^{-x^2}\cos ax\,dx.$ 7. $\displaystyle\int_1^\infty x^{\alpha-1}e^{-x}\,dx.$ 8. $\displaystyle\int_0^\infty e^{-x}\log(1+x)\,dx.$

9. $\displaystyle\int_0^\infty G(x)e^{-x}\,dx,$ $G(x)$, any polynomial.

10. The same integral, when $G(x)$ is replaced by any fraction, $R(x)$.

11. $\displaystyle\int_1^\infty x^{\alpha-1}e^{-x}\log x\,dx.$ 12. $\displaystyle\int_1^\infty x^{\alpha-1}e^{-x}(\log x)^n\,dx,$ $0 < n.$

Prove the following theorems. The integrand $f(x)$ is assumed continuous in the interval $c \leqq x < \infty$.

13. If a constant k greater than unity exists such that

$$x^k f(x) \quad \text{remains finite *} \qquad (k > 1)$$

then the integral (2) converges.

14. If $xf(x)$ remains numerically greater than some positive constant h for all values of x from a definite point on:

$$h < |xf(x)|, \qquad g \leqq x < \infty, \qquad \text{where} \quad 0 < h, \qquad c \leqq g,$$

the integral (2) diverges.

7. Alternating Integrals. Consider the integral

$$\int_0^\infty \frac{\sin x}{x} dx.$$

When x approaches 0, the integrand approaches 1, and the fact that the integrand is not defined for $x = 0$ is unimportant (foot-note, p. 466).

All of the foregoing tests fail to establish either the convergence or the divergence of this integral. There is, however, a simple and direct treatment, analogous to that used in the case of *alternating series*, *Introduction*, p. 398. Write

$$\int_0^x \frac{\sin x}{x} dx = \left(\int_0^\pi + \int_\pi^{2\pi} + \cdots + \int_{(m-1)\pi}^{m\pi} + \int_{m\pi}^{x}\right) \frac{\sin x}{x} dx, \quad m\pi \leqq x < (m+1)\pi.$$

The terms in the parenthesis are alternately positive and negative, as is seen by plotting the graph of the integrand,

$$y = \frac{\sin x}{x}.$$

Since the arches steadily flatten down, the numerical value of the *area bounded by an arch and the axis of x steadily* diminishes and approaches 0, and thus we see that the integral converges. The details are as follows. Compare

$$\int_{(m-1)\pi}^{m\pi} \frac{\sin x \, dx}{x} \qquad \text{with} \qquad \int_{m\pi}^{(m+1)\pi} \frac{\sin x \, dx}{x}.$$

* *i.e.* numerically less than some positive constant, G:

$$|x^k f(x)| < G, \qquad c \leqq x < \infty.$$

In the first integral, set

$$x = (m-1)\pi + t, \qquad \text{and let} \qquad x = m\pi + t$$

in the second integral. Thus we find

$$\pm \int_0^\pi \frac{\sin t\, dt}{(m-1)\pi + t} \qquad \text{and} \qquad \mp \int_0^\pi \frac{\sin t\, dt}{m\pi + t}.$$

Clearly, the numerical value of the second integral is less than that of the first. Moreover, the numerical value of either approaches 0 as its limit when $m = \infty$. Finally, if $m\pi \leqq x < (m+1)\pi$, the last integral approaches 0. Thus the given integral can be written as an alternating series, which satisfies all the conditions for convergence.

The Fresnel Integrals. In his investigations on light, Fresnel was led to the following integrals:

$$\int_0^\infty \frac{\sin x}{\sqrt{x}}\, dx, \qquad \int_0^\infty \frac{\cos x\, dx}{\sqrt{x}}.$$

The second of these involves the treatment of § 8 because the integrand becomes infinite at the lower limit of integration. The first can be shown to be convergent by the same reasoning as that set forth above.

Let us make the transformation

$$t = \sqrt{x}, \qquad x = t^2.$$

Then

$$\int_0^x \frac{\sin x\, dx}{\sqrt{x}} = 2 \int_0^t \sin t^2\, dt.$$

Since the variable on the left approaches a limit when $x = \infty$, the variable on the right must approach a limit when $t = \infty$, and thus we see that

$$\int_0^\infty \sin t^2\, dt$$

converges. And yet the integrand oscillates forever between $+1$ and -1. The explanation of the paradox lies in the fact that the graph of the function

$$y = \sin t^2,$$

though oscillating continually from $y = +1$ to $y = -1$, still is made up of arches whose bases are growing shorter and shorter. For, the curve crosses the axis when

$$t^2 = n\pi.$$

Thus the length of the base of an arch is $t_{n+1} - t_n$, where

$$t_{n+1}^2 = (n+1)\pi, \qquad t_n^2 = n\pi.$$

Hence $\quad t_{n+1}^2 - t_n^2 = \pi, \qquad$ or $\qquad t_{n+1} - t_n = \dfrac{\pi}{t_{n+1} + t_n} < \dfrac{\sqrt{\pi}}{2\sqrt{n}}.$

EXERCISES

1. Show that the integral

$$\int_0^\infty \frac{x \cos x}{a^2 + x^2}\, dx, \qquad a \neq 0,$$

converges.

2. Prove the theorem: If $\phi(x)$ is a positive monotonic decreasing function, continuous in the interval $c \leqq x < \infty$, and if $\lim\limits_{x=\infty} \phi(x) = 0$, then the integrals,

$$\int_c^\infty \phi(x) \sin x\, dx, \qquad \int_c^\infty \phi(x) \cos(ax + b)\, dx, \qquad a \neq 0,$$

converge.

3. Prove that the integral

$$\int_0^\infty \cos x^2\, dx$$

converges.

6. Infinite Integrands. The first example in § 3, (1) illustrates the case

(1) $$\int_a^b f(x)\, dx,$$

in which $f(x)$ is continuous at all points of the open interval

(2) $$a < x \leqq b,$$

but does not remain finite in the interval. The treatment of this case is parallel to that of the case just discussed at length, and the proofs are left to the student. First, the

Definition of Convergence. If

$$(3) \qquad \int_{a+\epsilon}^{b} f(x)\,dx$$

approaches a limit when ϵ approaches 0, the improper integral (1) is said to be *convergent*, and its *value* is defined as this limit:

$$(4) \qquad \int_{a}^{b} f(x)\,dx = \lim_{\epsilon \doteq 0} \int_{a+\epsilon}^{b} f(x)\,dx.$$

If the variable (3) approaches no limit, the integral (1) is said to be *divergent*, and no value is assigned to it.

It is readily shown by direct computation that

$$(5) \qquad \int_{a}^{b} \frac{C\,dx}{(x-a)^l}, \qquad 0 < l < 1,$$

converges (l, a constant). Also that the integral

$$(6) \qquad \int_{a}^{b} \frac{C\,dx}{x-a}, \qquad C \neq 0,$$

diverges.

On the integrals (5) and (6) are based the Comparison Tests and the Limit Tests for convergence and for divergence. It is assumed throughout that $f(x)$ is continuous at every point of the interval (2).

COMPARISON TEST FOR CONVERGENCE. *Let $\phi(x)$ be continuous and positive (or zero) in the interval* (2), *or in the left-hand part of that interval:*

$$0 \leqq \phi(x), \qquad a < x \leqq c \leqq b;$$

and let

$$\int_{a}^{c} \phi(x)\,dx$$

converge. If

$$0 \leqq f(x) \leqq \phi(x), \qquad a < x \leqq c,$$

then the integral (1) *converges.*

COMPARISON TEST FOR DIVERGENCE. *Let $\phi(x)$ be continuous and positive (or zero) as above:*

$$0 \leqq \phi(x), \qquad a < x \leqq c \leqq b,$$

and let

$$\int_a^b \phi(x)\,dx$$

diverge. If

$$f(x) \geqq \phi(x), \qquad a < x \leqq c,$$

then the integral (1) *diverges.*

Absolute Convergence. If

$$\int_a^b |f(x)|\,dx$$

converges, the integral (1) converges, and is said in this case to be *absolutely convergent.*

Limit Test for Convergence. *If there exists a positive constant l less than* 1 *such that $(x-a)^l f(x)$ approaches a limit when x approaches a:*

$$\lim_{x\to a}(x-a)^l f(x) \quad \textit{exists}, \qquad 0 < l < 1,$$

then the integral (1) *converges absolutely.*

The proof is given by the aid of the Comparison Test and the convergent integral (5).

Limit Test for Divergence. *If*

(*i*) $\lim_{x\to a}(x-a)f(x)$ *exists* — (denote it by L),

(*ii*) *and if the value of the limit is not zero,* $L \neq 0$;

then the integral (1) *diverges.*

Also, if $\lim_{x\to a}(x-a)f(x) = +\infty$ *or* $= -\infty$,

the integral (1) *diverges.*

There is no important class of alternating integrals, like those of § 7, to be considered here.

Example 1.
$$\int_0^1 x^{\alpha-1}e^{-x}\,dx.$$

When $\alpha \geqq 1$, the integral is a proper integral. When $\alpha = 0$, we have

$$\int_0^1 \frac{e^{-x}}{x}\,dx,$$

and the form of the integrand suggests divergence. In fact,

$$xf(x) = e^{-x},$$

and this function (*i*) approaches a limit (namely, 1) when x approaches 0; and (*ii*) the value of the limit is not 0. Hence the integral diverges. For any smaller value of α, the integrand is still larger, and hence we have divergence.

When $0 < \alpha < 1$, it is easy to see how to choose l, namely: $l = 1 - \alpha$. Then

$$x^l f(x) = e^{-x}, \qquad 0 < l < 1,$$

and since this function approaches a limit, the integral converges.

Example 2. $$\int_0^1 \left(\log \frac{1}{x}\right)^\gamma dx, \qquad 0 < \gamma.$$

Here, any value of l between 0 and 1 can be used to prove convergence:

$$x^l f(x) = x^l(-\log x)^\gamma = \left[\frac{-\log x}{x^{-l/\gamma}}\right]^\gamma.$$

It is readily shown by the usual method (Chap. X, § 2) that the function within the brackets approaches 0 when x approaches 0, and hence the whole expression approaches a limit (namely, 0).

9. Continuation. Consider the integral

$$\int_0^1 \frac{dx}{\sqrt{1-x}}.$$

Here, the integrand becomes infinite at the right-hand end of the interval.

The above is an example of the integral

(1) $$\int_a^b f(x)\,dx,$$

where $f(x)$ is continuous at every point of the open interval

(2) $$a \leq x < b.$$

This integral is defined to be *convergent* if

(3) $$\int_a^{b-\eta} f(x)\,dx$$

approaches a limit when η approaches 0, and its *value* is defined as this limit:

$$(4) \qquad \int_a^b f(x)\,dx = \lim_{\eta \doteq 0} \int_a^{b-\eta} f(x)\,dx.$$

If the integral (3) approaches no limit, the integral (1) is said to be *divergent*, and no value is assigned to it.

It is readily shown by direct computation that

$$(5) \qquad \int_a^b \frac{C\,dx}{(b-x)^l}, \qquad 0 < l < 1,$$

converges (l, a constant). Also, that the integral

$$(6) \qquad \int_a^b \frac{C\,dx}{b-x}, \qquad C \neq 0,$$

diverges.

On these integrals are based the Comparison Tests and the Limit Tests for convergence and for divergence.

The Comparison Tests for Convergence and Divergence are similar to those of § 8, the interval $a < x \leqq c$ now being replaced by the interval

$$c \leqq x < b, \qquad \text{where} \qquad a \leqq c < b.$$

The student should write these tests out in detail, and prove them.

Absolute Convergence is treated as in the earlier cases. Finally, the Limit Tests for Convergence and Divergence are similar to those of § 8, the variables on whose limits the tests turn now being

$$(b-x)^l f(x), \qquad 0 < l < 1$$

and

$$(b-x) f(x).$$

The student should also write out these tests in detail, and prove them.

Example. $\displaystyle\int_0^1 \frac{dx}{\sqrt{1-x^4}}$. Since $1 - x^4$ contains the simple linear factor $1 - x$, it is clear that $l = \frac{1}{2}$ will be a value of l, for which the Limit Test for Convergence will apply:

$$(1-x)^{\frac{1}{2}} f(x) = \frac{\sqrt{1-x}}{\sqrt{1-x^4}} = \frac{1}{\sqrt{(1+x)(1+x^2)}}.$$

Hence $\lim\limits_{x \doteq 1} (1-x)^{\frac{1}{2}} f(x)$ exists, and the integral converges.

10. Discontinuities within the Interval. The second example of § 3, (1) illustrates the case

(1) $$\int_a^b f(x)\,dx,$$

in which $f(x)$ is continuous at all but a finite number of points $\xi_1, \xi_2, \cdots, \xi_n$ of the interval $a \leqq x \leqq b$, or the interval $c \leqq x < \infty$.

Let $c_1, c_2, \cdots$ be points chosen, one in each of the sub-intervals into which the interval (a, b) is divided by the points $\xi_1, \xi_2, \cdots$. Consider the integral of $f(x)$ in each of the latter sub-intervals. In some, it may be a proper integral. If it converges in every one of these sub-intervals in which it is improper, then the integral (1) is said to be *convergent*, and its *value* is defined as the sum of the integrals, proper or improper, in these intervals. In all other cases, the integral (1) is *divergent*, and no value is assigned to it.

Thus the second integral of § 3, (1) diverges.

Example. The BETA FUNCTION, $B(m, n)$, is defined by the integral

$$B(m, n) = \int_0^1 x^{m-1}(1-x)^{n-1}\,dx,$$

when this integral converges. For what values of m and n will convergence take place?

Let c be a constant between 0 and 1. Then each of the integrals

$$\int_0^c x^{m-1}(1-x)^{n-1}\,dx, \qquad \int_c^1 x^{m-1}(1-x)^{n-1}\,dx$$

must either be a proper integral or a convergent improper integral.

It is easy to apply the foregoing tests to each of these integrals and thus to show that the Beta Integral converges for all points (m, n) lying within the first quadrant, and for no others.

EXERCISES

1. Carry through the details of the proof of convergence of the Beta Integral.

2. By making the change of variable, $t = 1 - x$, show that

$$B(m, n) = B(n, m).$$

3. **For what values of p and q will the integral**

$$\int_0^{\frac{\pi}{2}} \sin^{p-1} x \cos^{q-1} x \, dx$$

converge?

Are the following integrals convergent?

4. $\displaystyle\int_0^1 \frac{dx}{\sqrt{x}\log x}$.

5. $\displaystyle\int_0^{-\infty} e^x \log x^2 \, dx$.

6. $\displaystyle\int_{-\infty}^{\infty} \frac{e^{-x^2}\,dx}{\log x^2}$.

7. $\displaystyle\int_0^1 \frac{\log x\,dx}{1-x}$.

8. $\displaystyle\int_{-2}^{2} \frac{dx}{\sqrt[3]{-\log x^2}}$.

9. $\displaystyle\int_{-1}^{1} e^{-\frac{1}{x}}\,dx$.

11. The Gamma Function. Consider the integral

$$(1) \qquad \int_0^{\infty} x^{\alpha-1} e^{-x}\,dx.$$

This integral will converge if each of the integrals

$$(2) \qquad \int_0^1 x^{\alpha-1}e^{-x}\,dx, \qquad \int_1^{\infty} x^{\alpha-1}e^{-x}\,dx$$

which is improper, converges.

The second of these integrals converges for all values of α; § 6 Ex. 7. The first is a proper integral if $\alpha \geqq 1$. When $\alpha \leqq 0$, it diverges; but it converges if $0 < \alpha < 1$.

Hence the integral (1) converges for all positive values of α, and for no others. Its value is the Gamma Function,

$$(3) \qquad \Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1}e^{-x}\,dx, \qquad 0 < \alpha < \infty$$

The Difference Equation. The Gamma Function satisfies the difference equation:

$$(4) \qquad \Gamma(\alpha+1) = \alpha\Gamma(\alpha).$$

For, if we form the integral

$$\int_0^{\infty} x^{\alpha} e^{-x}\,dx$$

and then integrate by parts, we find:

$$\int_0^x x^a e^{-x}\,dx = -x^a e^{-x} + a\int_0^x x^{a-1}e^{-x}\,dx.$$

Now allow x to increase without limit. The left-hand side approaches $\Gamma(\alpha+1)$; the first term on the right approaches 0, and the second term approaches $\alpha\Gamma(\alpha)$, q. e. d.

The Factorial Function. The value of $\Gamma(1)$ has already been computed, § 3, Example:

$$\Gamma(1) = 1.$$

From the difference equation (4) we infer, on setting $\alpha = 1, 2, 3, \cdots, n+1$, successively, that

$$\Gamma(2) = 1\cdot 1, \quad \Gamma(3) = 2\Gamma(2) = 2\cdot 1 = 2!, \quad \Gamma(4) = 3\cdot 2\cdot 1 = 3!, \cdots$$

(5) $$\Gamma(n+1) = n!.$$

For this reason the Gamma Function is sometimes called the Factorial Function. It interpolates between the values of $n!$.

The Graph. We have:

(6) $$\Gamma(0^+) = +\infty.$$

For, $$\Gamma(\alpha) = \int_0^1 x^{\alpha-1}e^{-x}\,dx + \int_1^\infty x^{\alpha-1}e^{-x}\,dx > \int_0^1 x^{\alpha-1}e^{-x}\,dx.$$

Since $e^{-x} > e^{-1}$ in the interval in question,

$$\Gamma(\alpha) > \int_0^1 x^{\alpha-1}e^{-1}\,dx = \frac{1}{e\alpha},$$

and thus the truth of (6) is established.

That the function $\Gamma(\alpha)$, as defined by the integral (3), is continuous, is true; but the proof belongs to a later stage in analysis. Likewise, the fact that the derivative of the function is given by Leibniz's Rule:

(7) $$\Gamma'(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}\log x\,dx,$$

and that the second derivative is found by applying Leibniz's Rule again:

(8) $$\Gamma''(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}(\log x)^2\,dx.$$

These integrals can, however, be shown to converge, by our present methods. We observe, in particular, that the integrand in (8) is always positive, and hence

$$\Gamma''(\alpha) > 0, \qquad 0 < \alpha < \infty.$$

Thus the graph of the Gamma Function,

$$(9) \qquad y = \Gamma(\alpha), \qquad 0 < \alpha < \infty,$$

is always concave upward. Since it has the positive axis of y as an asymptote and goes through the points $(\alpha, y) = (n+1, n!)$, it is easy to plot the curve in character.

The Function $\Gamma(\alpha)$ *for* $\alpha < 0$. When $\alpha < 0$, the integral (1) diverges and so fails to define a function. We can, however, extend the definition by means of the difference equation, (4). Let α lie in the interval

$$-1 < \alpha < 0.$$

Then $\Gamma(\alpha+1)$ is defined for these values of α. And now we agree to define $\Gamma(\alpha)$ there by means of (4):

$$(10) \qquad \Gamma(\alpha) = \frac{\Gamma(\alpha+1)}{\alpha}.$$

Thus $\Gamma(\alpha)$ is negative in this interval. The graph (9) is continuous, and it evidently has the lines $\alpha = 0$ and $\alpha = -1$ as asymptotes. Moreover, it can be shown to be concave downward throughout the interval.

The process can be repeated. Let α lie in the interval

$$-2 < \alpha < -1,$$

and define $\Gamma(\alpha)$ by means of (10). The function is positive in this interval. The graph of (9) is continuous, and it has the lines $\alpha = -1$ and $\alpha = -2$ as asymptotes; it is concave upward. And so on indefinitely.

The student can now readily plot the curve in character. An accurately drawn graph is shown by Duval in the *Annals of Mathematics*, ser. 2, vol. 5 (1904), p. 65.

Tables of the Gamma Function. The function $\Gamma(\alpha)$ has been tabulated.* Because of the difference equation (4) it is sufficient to construct such a table for values of the argument from 1 to 2.

* Legendre, *Traité des fonctions elliptiques*, Paris, 1826, vol. II, p. 489.

Example. To determine $\Gamma(2.515)$ from Peirce's *Tables*, p. 140. Let $\alpha = 1.515$. Then

$$\Gamma(2.515) = 1.515 \times \Gamma(1.515) = 1.343.$$

Evaluation of $\Gamma(\frac{1}{2})$. If $\alpha = \frac{1}{2}$ in the integral (1), and if we make the change of variable, $t = \sqrt{x}$, we find:

$$\Gamma(\tfrac{1}{2}) = 2\int_0^\infty e^{-t^2}\,dt.$$

This latter integral has the value $\frac{1}{2}\sqrt{\pi}$; *cf.* § 13. Hence

(11) $$\Gamma(\tfrac{1}{2}) = \sqrt{\pi}.$$

EXERCISES

1. By means of the transformation $t = e^{-x}$, show that

$$\Gamma(\alpha) = \int_0^1 \left(\log\frac{1}{t}\right)^{\alpha-1} dt.$$

2. Prove: $\displaystyle\int_0^\infty x^n e^{-ax}\,dx = \frac{\Gamma(n+1)}{a^{n+1}}, \qquad 0 < a, \qquad -1 < n.$

Suggestion. Let $y = ax$.

3. Prove: $\displaystyle\int_0^1 x^m\left(\log\frac{1}{x}\right)^n dx = \frac{\Gamma(n+1)}{(m+1)^{n+1}}, \qquad -1 < n, \qquad -1 < m.$

Suggestion. Let $t = x^{m+1}$ in Question 1.

4. Test the integral of Question 1 directly for convergence and divergence.

5. Compute $\Gamma'(\alpha)$ and $\Gamma''(\alpha)$ by Leibniz's Rule from the integral of Question 1, and show that the resulting integrals converge.

6. Compute: $\displaystyle\int_0^1 \sqrt[3]{x\log(1/x)}\,dx$

7. Prove that

$$1\cdot 3\cdot 5\cdots(2n-1) = \frac{2^n\,\Gamma(n+\frac{1}{2})}{\sqrt{\pi}}.$$

8. Prove that

$$\frac{1\cdot 3\cdot 5\cdots(2n-1)}{2\cdot 4\cdot 6\cdots\ \ 2n} = \frac{\Gamma(n+\frac{1}{2})}{\sqrt{\pi}\,\Gamma(n+1)}.$$

12. The Beta Function. We have defined the Beta Function by the integral (§ 10)

(1) $$B(m, n) = \int_0^1 x^{m-1}(1-x)^{n-1}\,dx, \qquad 0 < n, \quad 0 < m,$$

and shown that

(2) $$B(m, n) = B(n, m).$$

If we make the change of variable in (1):

$$y = \frac{x}{1-x}, \qquad x = \frac{y}{1+y},$$

we find:

(3) $$B(m, n) = \int_0^\infty \frac{y^{m-1}\,dy}{(1+y)^{m+n}}.$$

It is possible to evaluate the Beta Function by means of the Gamma Function. The formal work is as follows. From § 11, Ex. 2, we have, on setting $n + 1 = m$:

$$\Gamma(m) = \int_0^\infty a^m x^{m-1} e^{-ax}\,dx.$$

Hence $$\Gamma(m)\,a^{n-1}e^{-a} = \int_0^\infty a^{m+n-1}x^{m-1}e^{-a(1+x)}\,dx.$$

Now, integrate each side of this equation with respect to a from 0 to ∞. If the iterated integral on the right were a proper integral, we could reverse the order of the integrations and thus obtain:

(4) $$\Gamma(m)\int_0^\infty a^{n-1}e^{-a}\,da = \int_0^\infty x^{m-1}\,dx \int_0^\infty a^{m+n-1}e^{-a(1+x)}\,da.$$

The value of the integral on the left is $\Gamma(n)$; that of the first integral to be computed on the right is, by § 11, Ex. 2:

$$\frac{\Gamma(m+n)}{(1+x)^{m+n}}.$$

Hence the right-hand side of (4) reduces, by the aid of (3), to

$$\Gamma(m+n)\int_0^\infty \frac{x^{m-1}\,dx}{(1+x)^{m+n}} = \Gamma(m+n)\,B(m, n).$$

Thus we find:

(5) $$B(m, n) = \frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)}.$$

The above reversal of the order of integrations is a question of double limits and requires proof. The proof, however, belongs to a later stage in analysis.

EXERCISES

Prove the following equations to be correct.

1. $$\int_0^1 \frac{x^{2n}\,dx}{\sqrt{1-x^2}} = \frac{\sqrt{\pi}\,\Gamma(n+\frac{1}{2})}{2\,\Gamma(n+1)}.$$

2. $$\int_0^{\frac{\pi}{2}} \sin^n x \cos^m x\,dx = \frac{\Gamma\left(\frac{m+1}{2}\right)\Gamma\left(\frac{n+1}{2}\right)}{2\,\Gamma\left(\frac{m+n}{2}+1\right)}.$$

3. $$\int_0^1 \frac{dx}{\sqrt{1-x^n}} = \frac{\sqrt{\pi}\,\Gamma\left(\frac{1}{n}\right)}{n\,\Gamma\left(\frac{1}{n}+\frac{1}{2}\right)}.$$

4. $$\int_0^1 x^m(1-x^n)^p\,dx = \frac{\Gamma(p+1)\,\Gamma\left(\frac{m+1}{n}\right)}{n\,\Gamma\left(p+1+\frac{m+1}{n}\right)}.$$

13. Improper Double Integrals. Let $f(x, y)$ be continuous at every point of the region

$$S: \qquad 0 \leq x < \infty, \qquad 0 \leq y < \infty.$$

The improper double integral

(1) $$\iint_S f(x, y)\,dS$$

is said to *converge* if the corresponding integral, extended over a variable finite region Σ lying in the first quadrant, approaches a limit, no matter how Σ expands, provided merely that an arbitrary finite region τ lying in S ultimately is included in all the regions Σ from a definite region $\bar{\Sigma}$ on. In case the integral (1) converges, the limits corresponding to the various possible choices of the variable regions Σ are all equal, and this is the *value* assigned to the integral (1).

Thus when (1) converges, it necessarily converges *absolutely*, *i.e.*

$$\iint_S |f(x, y)|\, dS$$

converges.

If, in particular, $f(x, y)$ does not change sign in S, it is obviously necessary and sufficient for the convergence of the integral (1) that it converge for one special set of regions Σ chosen as above.

The extension of the definition of convergence to other open regions S, whether finite or infinite, in which $f(x, y)$ is continuous, is obvious.

Consider in particular

$$\iint_S e^{-x^2-y^2}\, dS, \tag{2}$$

extended over the first quadrant. Let the region Σ be chosen as the quadrant of a circle, $0 \leqq r \leqq R$. Then

$$\iint_\Sigma e^{-x^2-y^2}\, dS = \int_0^{\frac{\pi}{2}} d\theta \int_0^R r e^{-r^2}\, dr = \frac{\pi}{4}(1 - e^{-R^2}).$$

Hence the integral (2) converges and has the value $\pi/4$.

Next, choose as the region Σ the square

$$0 \leqq x \leqq A, \qquad 0 \leqq y \leqq A.$$

We have:

$$\iint_\Sigma e^{-x^2-y^2}\, dS = \int_0^A e^{-y^2}\, dy \int_0^A e^{-x^2}\, dx = \left(\int_0^A e^{-x^2}\, dx\right)^2.$$

Now let A become infinite. Thus we infer that

$$\left(\int_0^\infty e^{-x^2}\, dx\right)^2 = \frac{\pi}{4}, \qquad \text{or} \qquad \int_0^\infty e^{-x^2}\, dx = \frac{\sqrt{\pi}}{2}, \tag{3}$$

and this latter important integral is evaluated.

14. Evaluation of Definite Integrals by Differentiation. It is sometimes possible to determine the function defined by a definite integral by differentiation. Only the formal work can be given here, since the justification of Leibniz's Rule in the case of improper integrals requires a more extended study of analysis.

Example 1. Consider the integral

(1) $$u = \int_0^\infty e^{-x^2} \cos \alpha x \, dx.$$

It converges absolutely for all values of α, since the integral

(2) $$\int_0^\infty e^{-x^2} dx = \frac{\sqrt{\pi}}{2}$$

converges; § 13. Differentiate (1):

(3) $$\frac{du}{d\alpha} = -\int_0^\infty x e^{-x^2} \sin \alpha x \, dx.$$

This integral is also seen to converge. Transform it by integration by parts, taking $xe^{-x^2} dx$ as one of the factors. Thus we find:

$$\int_0^\infty x e^{-x^2} \sin \alpha x \, dx = \frac{\alpha}{2} \int_0^\infty e^{-x^2} \cos \alpha x \, dx.$$

Hence (3) becomes:

$$\frac{du}{d\alpha} = -\frac{\alpha u}{2}.$$

Integrating, we have:

$$\log u = -\tfrac{1}{4}\alpha^2 + C, \qquad \text{or} \qquad u = ke^{-\frac{1}{4}\alpha^2}.$$

To determine the constant of integration, set $\alpha = 0$. From (2) we have $u = \sqrt{\pi}/2 = k$, and the final evaluation of the given integral is

(4) $$\int_0^\infty e^{-x^2} \cos \alpha x \, dx = \frac{\sqrt{\pi}}{2} e^{-\frac{1}{4}\alpha^2}.$$

Example 2. Let

(5) $$u = \int_0^\infty e^{-x^2 - \alpha^2/x^2} dx.$$

Here, $$\frac{du}{d\alpha} = -2\alpha \int_0^\infty x^{-2} e^{-x^2 - \alpha^2/x^2} dx.$$

And now it is a skillfully chosen substitution that leads to the result. Let $y = \alpha/x$ $\quad (0 < \alpha)$. Thus

$$\frac{du}{d\alpha}=-2u, \qquad u=Ce^{-2\alpha}, \qquad C=\frac{\sqrt{\pi}}{2};$$

(6) $$\int_0^\infty e^{-x^2-\alpha^2/x^2}\,dx=\tfrac{1}{2}\sqrt{\pi}e^{-2\alpha}, \qquad 0\leq\alpha;$$

or

(7) $$\int_0^\infty e^{-x^2-\alpha^2/x^2}\,dx=\tfrac{1}{2}\sqrt{\pi}e^{-2|\alpha|}, \qquad \text{(no restrictions on } \alpha).$$

15. Other Methods. Sometimes the form of the integrand suggests a special method.

Example 1. $\int_0^1 \frac{\log x\,dx}{1-x}$. Develop the factor $1/(1-x)$ into a series:

$$\frac{1}{1-x}=1+x+x^2+\cdots.$$

Hence (*Tables*, No. 427),

$$\int_0^1\frac{\log x\,dx}{1-x}=\sum_{n=0}^{\infty}\int_0^1 x^n\log x\,dx=-\left(\frac{1}{1^2}+\frac{1}{2^2}+\frac{1}{3^2}+\cdots\right).$$

It happens that the value of this series is well known (*cf.* Pierpont, *Functions of a Complex Variable*, p. 289):

(1) $$\frac{\pi^2}{6}=\frac{1}{1^2}+\frac{1}{2^2}+\frac{1}{3^2}+\cdots.$$

Thus the value of the integral is obtained. The proof, however, is incomplete, since it remains to show that the series can be integrated term-by-term. This proof belongs to a later stage in analysis.

Example 2. $\int_0^\infty \frac{\sin x}{x}\,dx.$ If we make the change of variable $x=\alpha y$ (and then replace y by x), we find that

(2) $$u=\int_0^\infty\frac{\sin\alpha x}{x}\,dx$$

has the same value for all positive values of α. Now, write

$$(3) \qquad u = \int_0^\pi \frac{\sin \alpha x}{x}\,dx + \int_\pi^\infty \frac{\sin \alpha x}{x}\,dx.$$

Applying the method of integration by parts to the second integral, we find:

$$(4) \qquad \int_\pi^\infty \frac{\sin \alpha x}{x}\,dx = \frac{\cos \pi\alpha}{\pi\alpha} - \frac{1}{\alpha}\int_\pi^\infty \frac{\cos \alpha x}{x^2}\,dx.$$

Let α become infinite. The limit of the first term on the right is 0. The integral in the second term is numerically less than

$$\int_\pi^\infty \frac{dx}{x^2} = \frac{1}{\pi}.$$

Hence the second term also approaches 0. Thus the integral on the left of (4) approaches 0.

The limit $\lim\limits_{\alpha=\infty} \int_0^\pi \frac{\sin \alpha x}{x}\,dx$ is one we have not met, nor is it readily determined. But it is well known in the theory of Fourier's Series and is shown to have the value $\pi/2$.

Thus we have, finally:

$$\int_0^\infty \frac{\sin \alpha x}{x}\,dx = \begin{cases} \dfrac{\pi}{2} & \text{when } \alpha > 0, \\ 0 & \text{" } \alpha = 0, \\ -\dfrac{\pi}{2} & \text{" } \alpha < 0. \end{cases}$$

Contour Integration in the Complex Domain. By means of Cauchy's Integral Theorem, Chap. XX, § 14:

$$\int_C f(z)\,dz = 0,$$

many definite integrals in the domain of real quantities can be evaluated. The method is set forth in treatises on the Theory of Functions of a Complex Variable; cf Goursat's *Mathematical Analysis* (translation by Hedrick and Dunkel), vol. II, p. 98, and the author's *Funktionentheorie*, vol. I, Chap. 7, § 3.

Further References. Byerly's *Integral Calculus*, 2d ed., Chap. VIII, contains much that is suggestive, particularly §§ 91–94; but § 95 is not helpful. Schlömilch, *Compendium der höheren Analysis*, vol. II, 3d ed., contains a chapter on the Gamma Function which is fairly accessible for the beginner.

EXERCISES

1. $\displaystyle\int_0^1 \frac{\log x\,dx}{1+x} = -\frac{\pi^2}{12}.$

2. $\displaystyle\int_0^1 \frac{\log x\,dx}{1-x^2} = -\frac{\pi^2}{8}.$

3. $\displaystyle\int_0^1 \frac{1}{x}\log\frac{1+x}{1-x}\,dx = \frac{\pi^2}{4}.$

4. $\displaystyle\int_0^\infty \log\frac{e^x+1}{e^x-1}\,dx = \frac{\pi^2}{4}.$

5. $\displaystyle\int_0^\infty \frac{e^{-ax}\sin mx}{x}\,dx = \tan^{-1}\frac{m}{a}.$ Suggestion: Differentiate.

6. $\displaystyle\int_0^\infty \frac{\cos mx\,dx}{1+x^2} = \frac{\pi}{2}e^{-|m|}.$ Suggestion: $\displaystyle\frac{1}{1+x^2} = 2\int_0^\infty \alpha e^{-\alpha^2(1+x^2)}\,d\alpha.$

CHAPTER XX

COMPLEX NUMBERS AND THE THEORY OF FUNCTIONS

1. The Age of Fable. Imaginary numbers came into the science through the attempt to obtain a solution of the quadratic equation in all cases. The particular quadratic,

$$x^2 + 1 = 0,$$

obviously cannot be satisfied by any integer, fraction, or incommensurable number, whether positive or negative, and these were the only numbers known to the science when the calculus was invented.

If a pure quadratic be written in the form

$$x^2 = a,$$

where a denotes a positive number or zero, then a root of this equation is

$$x = \sqrt{a}.$$

If $a = -1$, we can still make the mark, $\sqrt{-1}$; but it has no meaning. For only those things have meaning in mathematics which have been defined.

This statement must not, however, be understood as barring the road to further definitions, and it was, in fact, this thought, — the possibility of *finding a meaning*, as it appeared to the mathematicians of the eighteenth century; of *extending the number system* by introducing a new *definition*, as we should say, — that, first vaguely, then ever more clearly, guided our predecessors in their quest for what they imagined to exist, but which they had not yet succeeded in grasping.*

The Formal Period. Let us follow the historical development of the idea. At first, $\sqrt{-1}$ was a mark,† which was written in algebraic

* Recall, for example, the words of Leibniz, quoted on p. 389.

† A "symbol," some would call it. But what is a *symbol*? Does a symbol

expressions where a letter, as a or x, might appear, and expressions containing it were transformed according to the five Formal Laws of Algebra:

$$A+B=B+A,$$
$$A+(B+C)=(A+B)+C,$$
$$AB=BA,$$
$$A(BC)=(AB)C,$$
$$A(B+C)=AB+AC.$$

Thus, in particular,

$$a+b\sqrt{-1}+(a'+b'\sqrt{-1})$$

was replaced at pleasure by

$$(a+a')+(b+b')\sqrt{-1}.$$

We have here the idea which underlies the later definition of addition for any two complex numbers, $a+b\sqrt{-1}$ and $a'+b'\sqrt{-1}$.

Furthermore

$$a\sqrt{-1} \qquad \text{and} \qquad \sqrt{-1}\,a$$

were considered as interchangeable. And two of these queer expressions, $a+b\sqrt{-1}$ and $a'+b'\sqrt{-1}$, were considered as equal (*i.e.* one could replace the other on any occasion) if and only if $a=a'$ and $b=b'$:

$$a+b\sqrt{-1}=a'+b'\sqrt{-1} \quad \text{if} \quad a=a' \quad \text{and} \quad b=b'.$$

In the case, however, of multiplication (*i.e.* in case two of these marks followed each other like two letters which are multiplied together, as ab, — for, of course, nothing is defined as yet, and so it is a question merely of an expression which looks as if it would like to be a product if it had a chance) a rule was adopted which went beyond the formal laws of algebra; for $\sqrt{-1}\sqrt{-1}$, it was agreed, should be replaceable at will by -1:

$$\sqrt{-1}\sqrt{-1}=-1.$$

With this understanding, it is possible to write, not merely

$$(a+b\sqrt{-1})(a'+b'\sqrt{-1})=aa'+(ab'+a'b)\sqrt{-1}+bb'\sqrt{-1}\sqrt{-1},$$

stand for something? If so, what is the thing for which $\sqrt{-1}$ stands? Or is a symbol an object having independent existence? If so, what is the object, $\sqrt{-1}$?

Such definitions of $\sqrt{-1}$ as "the indicated root of a negative quantity" leave the beginner mystified unless he has sufficient insight to discern that words are being formed which collectively have no meaning.

but to modify the result still further and set

$$(a+b\sqrt{-1})(a'+b'\sqrt{-1})=aa'-bb'+(ab'+a'b)\sqrt{-1}.$$

We have here the forerunner of the definition of multiplication.

Applications. This formal acceptance of $\sqrt{-1}$ opened the way to further developments of the greatest importance in the field of algebra and algebraic geometry on the one hand, and on the other, in analysis. For it was early surmised that every algebraic equation has a root in the domain of imaginaries. This granted, algebraic geometry mounts up from an insignificant mass of inequalities to a well-rounded science, homogeneous and having mathematical content. Thus a circle,

$$x^2+y^2=a^2,$$

is cut by a straight line,

$$Ax+By+C=0,$$

at most in two points, and the order of multiplicity of straight lines which do not meet the circle at all is as great as that of those which do — in the domain of reals. But in the domain of imaginaries, a straight line cuts the circle in general in just two points, and the exceptions (tangents and null-lines) form a manifold of lower order.

In analysis, the elementary transcendental functions ($\sin x$, $\cos x$, e^x, and their inverses) were well known in the domain of reals. What meaning is to be attached to them in the realm of imaginaries? This question presents itself in altogether concrete form when one seeks the solution of a linear differential equation with constant coefficients, and it is there that it is best treated. We ask the reader, therefore, to turn back to Chapter XIV, § 12 and follow the account of the heuristic considerations which finally led to the equation

$$e^{\phi\sqrt{-1}}=\cos\phi+\sqrt{-1}\sin\phi.$$

2. Geometrical Representations. (*a*) *Points of a Plane.* After the formal period had forecast the importance of $\sqrt{-1}$ and before modern ideas of rigor through definition had made progress, there was flashed upon the screen of mathematics a geometric picture which went far toward uniting the past of the imagination with the future of the reality. Simultaneously and independently, Gauss, Argand, and Wessel perceived that a point in the plane, whose coordinates are (x, y), can represent the imaginary number

$$z=x+y\sqrt{-1}.$$

In the word "represent" there is a *petitio*, for how can that which does not yet exist be represented? And yet the point exists, and the pair of numbers (x, y) exists. So it came about that mathematicians were easier in their minds concerning imaginaries, now that they had this other system of points in a plane to stand in the place of these mystical objects.*

(*b*) *Vectors.* A second geometric representation of the number $x + y\sqrt{-1}$ consists in the vector † drawn from the origin to the point (x, y). Let

$$\mathfrak{A} = a + bi, \qquad \mathfrak{B} = c + di,$$

(where from now on we shall replace $\sqrt{-1}$ by Euler's notation for it, namely, i) be any two complex ‡ numbers. They are defined as *equal*, $\mathfrak{A} = \mathfrak{B}$, if and only if $a = c$ and $b = d$. Their *sum*,

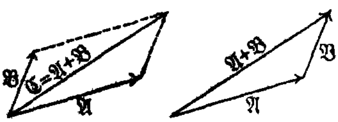


FIG. 117

$$\mathfrak{A} + \mathfrak{B} = \mathfrak{C},$$

is defined as the number

$$\mathfrak{C} = a + c + (b + d)i.$$

The vector which represents $\mathfrak{C}$ is constructed geometrically by the law of the parallelogramme of forces, and the construction is known as *vector addition*. Furthermore

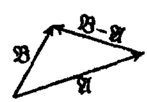


FIG. 118

$$\mathfrak{B} - \mathfrak{A} = c - a + (d - b)i.$$

Polar Coordinates. A complex number, $z = x + yi$, can be written in the polar form:

$$x + yi = r(\cos\phi + i \sin\phi).$$

* If we may anticipate the historical account, it remained only to recognize that the essence of a number-system is (*i*) a class of objects, and (*ii*) the postulated laws connecting them. In this sense we may say, if we like, that the points of the plane *are* the numbers. This thought, although formulated explicitly much later, was in the minds of mathematicians as soon as the geometric representation became known.

† The reader should now turn to the chapter on *Vector Analysis* and read the first paragraph.

‡ The term *imaginaries*, emphasizing the mystical, gradually gave way to the term *complex numbers*, emphasizing the dependence of $a + bi$ on the complex of real numbers, (a, b). We still speak, however, of i as the *imaginary unit*, and of ai, where a is real, as a *pure imaginary*.

Here, r is the *absolute value* of z:

$$|z| = r = \sqrt{x^2 + y^2}.$$

ϕ is called the *angle* of z, and is written:* $\phi = \text{arc}\, z$.

Multiplication. The *product* of two complex numbers, $z_1 = x_1 + y_1 i$ and $z_2 = x_2 + y_2 i$, is defined as suggested by the formal work in § 1:

$$z_1 z_2 = x_1 x_2 - y_1 y_2 + (x_1 y_2 + x_2 y_1) i.$$

When z_1 and z_2 are written in polar form,

$$z_1 = r_1(\cos\phi_1 + i \sin\phi_1), \qquad z_2 = r_2(\cos\phi_2 + i\sin\phi_2),$$

the product becomes

$$r_1 r_2\{\cos\phi_1\cos\phi_2 - \sin\phi_1\sin\phi_2 + (\sin\phi_1\cos\phi_2 + \cos\phi_1\sin\phi_2) i\}.$$

Hence

$$z_1 z_2 = r_1 r_2(\cos\overline{\phi_1 + \phi_2} + i\sin\overline{\phi_1 + \phi_2}).$$

We are thus led to the following rule: — *To multiply two complex quantities, multiply their absolute values and add their angles.*

Geometric Interpretation. The vector which represents the product, $z_1 z_2$, can be constructed geometrically as follows. Consider first the triangle whose vertices are at the points $0, 1, z_1$. Draw a second triangle similar to this one, in which the line joining 0 with the point z_2 corresponds to the side from 0 to 1. Then the other side which emanates from 0 will make an angle $\phi_1 + \phi_2$ with the positive axis of reals.† Denote its length by R. From the similarity of the triangles,

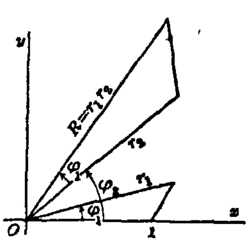


FIG. 119

$$\frac{R}{r_2} = \frac{r_1}{1}, \qquad \text{or} \qquad R = r_1 r_2.$$

Hence the third vertex of the triangle is at the point $z_1 z_2$.

The geometric construction emphasizes the important fact that, while addition is a process defined in terms of the vectors alone, irrespective of the coordinate axes and the unit of length, multipli-

* r is sometimes called the *modulus*, and ϕ, the *amplitude* or *argument*. But these words have so many other meanings in mathematics that they are not distinctive for the present purpose.

† The axis of x is called the *axis of reals*, and the axis of y, the *axis of pure imaginaries*.

cation bears jointly on the given vectors and the choice of the vector which represents the number 1.

Division. Division is defined as the inverse of multiplication:

$$w = \frac{z_1}{z_2} \qquad \text{if} \qquad wz_2 = z_1 \qquad \text{or} \quad z_2 w = z_1.$$

Thus $$\frac{z_1}{z_2} = \frac{r_1}{r_2}\left\{\cos \overline{\phi_1 - \phi_2} + i \sin \overline{\phi_1 - \phi_2}\right\}.$$

Division is always possible and unique except when the divisor, z_2, is 0, and then division is not defined.

The Formal Laws. Addition and multiplication obey the five Formal Laws of Algebra, § 1. In particular, the Commutative Law holds for multiplication as well as for addition:

$$z_1 z_2 = z_2 z_1.$$

Furthermore, as in the ordinary algebra of real numbers, a product vanishes when and only when one of its factors vanishes.

Definitions. If $a + bi$ is any complex quantity, then $a - bi$ is called the *conjugate imaginary*, or is said simply to be *conjugate* to it. A real quantity coincides with its conjugate.

We furthermore lay down the definition:

(*A*) $$e^{\phi i} = \cos \phi + i \sin \phi.$$

Thus $$e^{x+yi} = e^x (\cos y + i \sin y).$$

And again, if $z = x + yi$, then z can also be written as $re^{\phi i}$.

By the *unit circle* is meant the circle

$$x^2 + y^2 = 1.$$

EXERCISES

1. Show that $z^2 = x^2 - y^2 + 2xyi$, and find z^3.

2. Express $$\frac{1}{a + bi} \qquad \text{and} \qquad \frac{c + di}{a + bi}$$

in the reduced form of a complex number, $A + Bi$.

3. Prove that the sum of two conjugate imaginaries is real, and that their difference is a pure imaginary. Prove that their product is real. When will their quotient be real?

4. If $G(x)$ is a polynomial with real coefficients, and if z and $\bar{z}$ are conjugate imaginaries, prove that $G(z)$ and $G(\bar{z})$ are conjugate imaginaries.

5. If $R(x)$ is any rational function with real coefficients, and if z and $\bar{z}$ are conjugate imaginaries, prove that $R(z)$ and $R(\bar{z})$ are conjugate imaginaries.

6. Prove the theorem: if $|z| > 1$, the point which represents $1/z$ is obtained as follows. Draw the tangents from z to the unit circle. The intersection of the chord joining them, with the line from 0 to z, represents the conjugate imaginary of $1/z$.

3. Inequalities. If $\mathfrak{A}$ and $\mathfrak{B}$ be any two complex numbers, then

(1) $$|\mathfrak{A} + \mathfrak{B}| \leqq |\mathfrak{A}| + |\mathfrak{B}|.$$

For, any side of a triangle is less than the sum of the other two sides; *cf.* Fig. 117, § 2. Hence, for a true triangle, only the sign of inequality can hold. But if $\mathfrak{A}$ and $\mathfrak{B}$ are collinear, *i.e.* are parallel vectors, the sign of equality may hold. From (1) it follows generally that

(2) $$|A_1 + A_2 + \cdots + A_n| \leqq |A_1| + |A_2| + \cdots + |A_n|.$$

EXERCISES

1. Prove that
$$|[|\mathfrak{A}| - |\mathfrak{B}|]| \leqq |\mathfrak{A} - \mathfrak{B}|.$$

2. Prove that
$$\frac{1}{\sqrt{2}}[|a| + |b|] \leqq |a + bi| \leqq |a| + |b|.$$

4. Powers and Roots. If
$$A = \mathsf{A}(\cos\alpha + i\sin\alpha) = \mathsf{A}e^{\alpha i}$$
be any complex number, then
$$A^2 = \mathsf{A}^2(\cos 2\alpha + i\sin 2\alpha) = \mathsf{A}^2 e^{2\alpha i},$$
$$A^3 = \mathsf{A}^3(\cos 3\alpha + i\sin 3\alpha) = \mathsf{A}^3 e^{3\alpha i},$$
$$\cdots\cdots\cdots$$
$$A^n = \mathsf{A}^n(\cos n\alpha + i\sin n\alpha) = \mathsf{A}^n e^{n\alpha i}.$$

By an *n-th root* of A (n, a positive integer) we mean a number $z = re^{\phi i}$ such that

(1) $$z^n = A.$$

Hence $$r^n e^{n\phi i} = Ae^{\alpha i}.$$

The complex numbers which stand on the two sides of this equation, being equal, must have the same absolute values, and hence

$$r^n = A, \qquad r = \sqrt[n]{A}.$$

Their angles, however, need not be equal. It is necessary and sufficient that they differ from each other by an integral multiple of 2π or $360°$:

$$n\phi = \alpha, \quad \alpha + 2\pi, \quad \alpha + 4\pi, \quad \cdots,$$
$$\alpha - 2\pi, \quad \alpha - 4\pi, \quad \cdots.$$

In particular, we may take $0 \leqq \alpha < 2\pi$, and then we find n values of ϕ leading to distinct points of the plane, namely:

$$\phi = \frac{\alpha}{n}, \quad \frac{\alpha + 2\pi}{n}, \quad \frac{\alpha + 4\pi}{n}, \quad \cdots, \quad \frac{\alpha + 2k\pi}{n}, \quad \cdots, \quad \frac{\alpha + 2(n-1)\pi}{n}.$$

All other values of ϕ lead to one of these n points.

Hence it appears that the n roots of equation (1) are given by the formula:

(2) $$z_k = A^{\frac{1}{n}} e^{\frac{(\alpha + 2k\pi)i}{n}}, \qquad k = 0, 1, \cdots, n-1.$$

Discussion. Consider first the *roots of unity*:

$$z^n = 1.$$

Here, $A = 1$, $\alpha = 0$, and the roots lie on the unit circle at the points

$$1, \quad e^{\frac{2\pi}{n}i}, \quad e^{\frac{4\pi}{n}i}, \cdots, \quad e^{\frac{2(n-1)\pi}{n}i},$$

thus forming the vertices of a regular inscribed polygon of n sides.

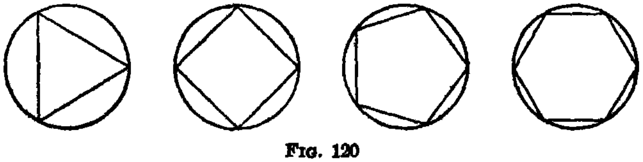


FIG. 120

It is now easy to interpret geometrically the general case. Draw first the circle $r = A^{\frac{1}{n}}$, and mark on it a point, P, such that the radius drawn to it makes an angle α/n with the positive axis of reals. Inscribe a regular polygon of n sides with one vertex at P. Then the vertices of this polygon represent the n roots of A.

EXERCISES

1. Determine all the roots of the equation

$$z^5 = 2 + 4i$$

and plot them, showing the corresponding pentagon.

2. Find the cube roots of -1 and plot them, showing the corresponding equilateral triangle. Similarly for the fourth, fifth, and sixth roots of -1.

3. Write the polynomial

$$x^4 + a^4$$

(*i*) as the product of its linear factors; (*ii*) as the product of its real quadratic factors.

4. The same for $x^6 + a^6$.

5. The sum of the n n-th roots of unity is zero. Give both an algebraic and a geometric proof.

6. Show that, if ω is an n-th root of unity, all the integral powers of ω are also roots of unity; and that, if ω be properly chosen, the n roots can be written:

$$\omega, \quad \omega^2, \quad \omega^3, \quad \cdots, \quad \omega^{n-1}, \quad \omega^n = 1.$$

One such choice for ω is $e^{\frac{2\pi i}{n}}$.

7. Solve the quadratic equation

$$z^2 = 1 + i$$

by writing z^2 in the form $x^2 - y^2 + 2xyi$ and then equating reals and pure imaginaries on the two sides of the equation. Plot the vectors, drawn from 0, which represent the roots, and show how they are related to the vector which represents $1 + i$.

Generalize for

$$z^2 = a + bi.$$

8. If n complanar forces, acting at a point, are represented by the complex numbers $z_1, z_2, \cdots, z_n$, show that a necessary and sufficient condition that these forces be in equilibrium is:

$$z_1 + z_2 + \cdots + z_n = 0.$$

5. The Function e^z. We have already defined $e^{\phi i}$ and e^z, where $z = x + yi$, in § 1:

(1) $$e^z = e^{x+yi} = e^x(\cos y + i \sin y).$$

But what justification, what reason, is there for this definition beyond the mere formal work of § 1?

The same question presented itself in elementary algebra when we passed from $a^n = a \cdot a \cdot \cdots \cdot a$ (n times), where n is a natural number, to the new and extended meaning of a^n, where n is a fraction or a negative rational number, and finally to the completed definition of a^x, where x is any real number, integral, fractional, or incommensurable. It was the *formal law*

(2) $$a^n a^m = a^{n+m}$$

that guided us then, and it is that law now which expresses the essential property of the exponential function. This law is known as the *Addition Theorem* for the exponential function and can be expressed in general terms as follows:

(3) $$f(x+y) = f(x)f(y),$$

where x and y are any two numbers — real, in the case of elementary algebra, complex for our proposed extension (1).

Does our new function (1) measure up to this standard? Is it true that

(4) $$e^{z_1+z_2} = e^{z_1}e^{z_2} \quad ?$$

Ask the definition. That is the only source of an answer.

$$\begin{aligned} e^{z_1}e^{z_2} &= e^{x_1}(\cos y_1 + i \sin y_1)\, e^{x_2}(\cos y_2 + i \sin y_2) \\ &= e^{x_1+x_2}(\cos \overline{y_1+y_2} + i \sin \overline{y_1+y_2}). \end{aligned}$$

But this last expression is precisely $e^{z_1+z_2}$, and the new function has stood the test.

Moreover, just as e^x is uniquely defined and is continuous for all real values of x, so e^z is single-valued and continuous for all complex values of z.

But is this all that we could ask for? Two further properties of the real function e^x are:

(5) (a) $\dfrac{de^x}{dx} = e^x$; (b) $e^x = 1 + x + \dfrac{x^2}{2!} + \cdots$.

That e^z also has these properties, will be shown in §§ 12 and 16.

Over and above that which could be expected we find, however, an entirely new property: — *The function e^z is periodic in the imaginary domain:*

$$e^{z+2\pi i} = e^z.$$

It has the imaginary period $2\pi i$.

EXERCISES

1. Show that $e^{\pi i} = -1, \quad e^{\frac{\pi i}{2}} = i.$

2. Compute the value of e^{1+i} in the form $a + bi$, determining a and b to four places of decimals.

3. Show that

$$|e^z| = e^x \quad \text{and} \quad |e^{\phi i}| = 1,$$

where $z = x + yi$ and ϕ is real.

6. The Function $\log z$. In the domain of reals the functions e^x and $\log x$ are inverses, the one of the other:

$$y = \log x \quad \text{if} \quad e^y = x.$$

In the domain of complex quantities we will adopt the same definition and say:

$$w = \log z \quad \text{if} \quad e^w = z.$$

Let $\qquad w = u + vi, \qquad z = r(\cos\phi + i\sin\phi).$

Then $\log z$ is given by the equation:

$$e^{u+vi} = r(\cos\phi + i\sin\phi).$$

Since the complex number which stands on the left of this equation is the same as the complex number on the right, it follows (*i*) that their absolute values must be equal; (*ii*) that their angles must be equal, save as to a multiple of 2π:

(*i*) $\qquad e^u = r, \qquad u = \log r;$

(*ii*) $\qquad v = \phi + 2k\pi.$

Hence

(1) $$\log z = \log r + \phi i,$$

where ϕ may be *any one* of the determinations of the angle of z.

Thus it appears that, while the real function $\log x$ was *single-valued* and defined for *positive values* of x only, the extended function is defined for *all complex values* of the argument, except 0, and is *infinitely multiple-valued.*

In particular, a negative number, $-a$, now has a logarithm:

$$\log(-a) = \log a + (2k+1)\pi i.$$

Thus $\qquad \log(-1) = \pi i, \quad \text{or} \quad -\pi i, \quad \text{or} \quad 3\pi i, \quad \text{etc.}$

Moreover, a positive number has an infinite number of logarithms

Thus $$\log 1 = 0, \quad \pm 2\pi i, \quad \pm 4\pi i, \cdots.$$

We note that

$$e^{\log A} = A, \qquad \log e^B = B + 2k\pi i,$$

where A is any complex number $\neq 0$, and B is wholly arbitrary.

Formula (1) is sometimes written in the Cartesian form:

$$\text{(2)} \qquad \log z = \tfrac{1}{2}\log(x^2 + y^2) + i\tan^{-1}\frac{y}{x}.$$

But this is wrong, since not every value of $\tan^{-1} y/x$ is admissible. Thus, for example, this formula would make

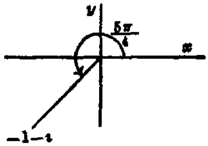


FIG. 121

$$\log(-1 - i) = \tfrac{1}{2}\log 2 + \frac{\pi}{4}i + k\pi i,$$

where k is any integer, whereas in fact k can here take on only *odd* values. For, *one* determination is

$$\log(-1 - i) = \tfrac{1}{2}\log 2 + \tfrac{5}{4}\pi i,$$

and the others differ from it by multiples, not of πi, but of $2\pi i$.

No confusion can arise if the value of ϕ in (1) is read off from the figure. There, ϕ is the angle, measured in radians, which belongs to z — *any* value of this angle.

The Functional Relations:

$$\text{(3)} \qquad \log z_1 + \log z_2 = \log(z_1 z_2);$$

$$\text{(4)} \qquad \log z^n = n \log z.$$

The first equation holds for any two complex numbers z_1, z_2, both different from 0, where any two of the three logarithms may be chosen arbitrarily among their possible values, and the third then suitably determined. The second relation depends on the extension of the definition of the function z^n; *cf. infra.*

The Generalized Power, A^B. Let A and B be any two complex numbers, provided $A \neq 0$. Then we define

$$\text{(5)} \qquad A^B = e^{B\log A},$$

where $\log A$ has any one of its possible values. Thus

$$i^i = e^{-\frac{\pi}{2}}, \quad e^{\frac{3\pi}{2}}, \quad \cdots, \quad e^{-\frac{\pi}{2} + 2k\pi}.$$

It would now appear as if there were a conflict with the former definition of e^z, since now

$$e^z = e^{z(\log e)} = e^z,\ e^{z(1+2\pi i)},\quad \cdots.$$

We cut the knot, however, by restricting the notation e^z to its meaning as given by § 5, (1).

The Functional Relation (4) is seen to be valid, where the logarithm on the right may be chosen at pleasure, and then both z^n and $\log z^n$ are in general uniquely determined.

EXERCISES

1. Find all the values of $\log 10$.

   Ans. $2.30259 + 2k\pi i,\quad k = 0,\ \pm 1,\ \pm 2, \cdots.$

2. Find all the values of $\log(2 - 3i)$.

   Ans. $1.2825 - 0.9828i + 2k\pi i.$

3. Find all the values of $\log(-5 - 6i)$.

4. Determine all the values of π^π.

5. Compute the values of $(1 + i)^i$.

   Ans. $(535.5)^k(0.4288 + 0.1548i),\quad k = 0,\ \pm 1,\ \pm 2, \cdots.$

6. Show that

$$A^B A^C = A^{B+C}, \qquad A \neq 0,$$

where any one of the three quantities A^B, A^C, A^{B+C} may be chosen arbitrarily among its possible values, but in general neither of the remaining two can be so chosen.

State precisely the latitude of choice in the relation

$$5^{\sqrt{2}}5^{\sqrt{3}} = 5^{\sqrt{2}+\sqrt{3}}.$$

7. The Functions $\sin z$, $\cos z$, $\tan z$, etc. From the equation of definition, § 1:

$$e^{\phi i} = \cos\phi + i\sin\phi, \tag{1}$$

where ϕ is real, follows that

$$e^{-\phi i} = \cos\phi - i\sin\phi.$$

Hence

$$\sin\phi = \frac{e^{\phi i} - e^{-\phi i}}{2i}, \qquad \cos\phi = \frac{e^{\phi i} + e^{-\phi i}}{2}. \tag{2}$$

Now that e^z is defined for all complex values of the argument, the right-hand sides of these equations have a meaning when ϕ is complex. What more natural than to take these extensions as the definition of the functions $\sin z$, $\cos z$ for complex values of the argument, and see what happens?

$$(3) \qquad \sin z = \frac{e^{iz} - e^{-iz}}{2i}, \qquad \cos z = \frac{e^{iz} + e^{-iz}}{2}.$$

The essential question is: What are the most important *analytical* properties of the trigonometric functions, $\sin x$, $\cos x$? Again, the answer is flashed back to us in the form of the *Addition Theorem* for these functions:

$$(4) \qquad \begin{cases} \sin(z_1 + z_2) = \sin z_1 \cos z_2 + \cos z_1 \sin z_2, \\ \cos(z_1 + z_2) = \cos z_1 \cos z_2 - \sin z_1 \sin z_2. \end{cases}$$

Unless the extended functions satisfy these relations, equations (3) are but a hollow nut. And so it is with curiosity that we compute the right-hand side of the first of these equations from the definition (3). On reducing the result by means of (4), § 5, we see that the equation is, indeed, true. And likewise for the second equation.

Moreover, the extended functions are single-valued and continuous for all values of the complex variable, z.

Is this all that we could desire? The real functions $\sin x$, $\cos x$ have derivatives, given by the familiar formulas, and the functions satisfy a linear differential equation of the second order, which in turn dominates these functions completely. And, finally, there are the power series expansions. Will these properties persist? The answer is most satisfactory. We shall show later that

$$(5) \qquad \frac{d \sin z}{dz} = \cos z, \qquad \frac{d \cos z}{dz} = -\sin z;$$

thus the functions $\sin z$, $\cos z$ are solutions of the linear differential equation of the second order:

$$(6) \qquad \frac{d^2 w}{dz^2} + w = 0;$$

and, finally, that

$$(7) \qquad \begin{cases} \sin z = z - \dfrac{z^3}{3!} + \dfrac{z^5}{5!} - \cdots, \\[2ex] \cos z = 1 - \dfrac{z^2}{2!} + \dfrac{z^4}{4!} - \cdots. \end{cases}$$

The other trigonometric functions are defined in terms of $\sin z$, $\cos z$:

$$(8) \qquad \tan z = \frac{\sin z}{\cos z} = \frac{1}{i}\frac{e^{iz} - e^{-iz}}{e^{iz} + e^{-iz}};$$

(9) $\cot z = \dfrac{1}{\tan z}, \qquad \sec z = \dfrac{1}{\cos z}, \qquad \csc z = \dfrac{1}{\sin z}.$

In the domain of reals there is no relation between the trigonometric functions on the one hand, and the exponential function on the other. In the domain of imaginaries, we have here before us only a single class of functions.

EXERCISES

1. Compute $\sin(1+i)$. *Ans.* $1.299 + 0.635\,i$.

2. Express e^z in terms of $\tan z$. *Ans.* $e^z = \dfrac{1 - i\tan\frac{iz}{2}}{1 + i\tan\frac{iz}{2}}$.

8. The Inverse Trigonometric Functions, $\sin^{-1} z$, $\tan^{-1} z$, etc. The anti-sine is defined by the equations:

(1) $$w = \sin^{-1} z \quad \text{if} \quad \sin w = z.$$

For the determination of w we have, then:

$$\frac{e^{iw} - e^{-iw}}{2i} = z.$$

Hence

$$e^{2iw} - 2iz\,e^{iw} - 1 = 0,$$

$$e^{iw} = iz \pm \sqrt{1-z^2}, \qquad w = \frac{1}{i}\log(iz \pm \sqrt{1-z^2}).$$

The final result can be written in a variety of ways (*cf.* Peirce's *Tables*, p. 79, No. 643). Since

$$-\log(iz \pm \sqrt{1-z^2}) = \log\frac{1}{iz \pm \sqrt{1-z^2}}$$

and

$$\frac{1}{iz \pm \sqrt{1-z^2}} = \pm\sqrt{1-z^2} - iz,$$

we may write:

(2) $$\sin^{-1} z = i\log[\sqrt{1-z^2} - iz],$$

where it is understood that *both* values of the radical must be observed.

The anti-cosine and the anti-tangent are defined and computed in a similar fashion:

$$(3)\qquad \begin{cases} w = \cos^{-1} z & \text{if} \quad \cos w = z, \\ \cos^{-1} z = i \log(z + i\sqrt{1 - z^2}). \end{cases}$$

$$(4)\qquad \begin{cases} w = \tan^{-1} z & \text{if} \quad \tan w = z, \\ \tan^{-1} z = \dfrac{i}{2} \log \dfrac{i + z}{i - z}. \end{cases}$$

As was to be expected, the inverse trigonometric functions and the logarithm form, in the complex domain, but a single class of functions.

EXERCISES

Compute the following, determining *all* values:

1. $\sin^{-1} 2$. 2. $\cos^{-1} 5$. 3. $\tan^{-1}(1 + i)$.

4. Show that the two formulas 49 and 50 given in Peirce's *Tables* for

$$\int \frac{dx}{a + bx^2}$$

are identical in the domain of imaginaries.

9. The Hyperbolic Functions. Certain functions analogous to the trigonometric functions, called the *hyperbolic functions*, have recently come into general use. They go back, however, to Riccati (1757) and are defined as follows:

$$\sinh x = \frac{e^x - e^{-x}}{2};$$

$$\cosh x = \frac{e^x + e^{-x}}{2};$$

$$\tanh x = \frac{\sinh x}{\cosh x},$$

etc. (read "hyperbolic sine of x," etc.). An abbreviated notation for $\sinh x$, $\cosh x$, $\tanh x$, is $\operatorname{sh} x$, $\operatorname{ch} x$, $\operatorname{th} x$. The graphs of these functions are shown in Fig. 122.* The functions satisfy the following relations, $\operatorname{sh} x$ and $\operatorname{th} x$ being odd functions, $\operatorname{ch} x$ an even function:

$$\operatorname{sh}(-x) = -\operatorname{sh} x, \quad \operatorname{ch}(-x) = \operatorname{ch} x, \quad \operatorname{th}(-x) = -\operatorname{th} x.$$

* The graph of the function $\operatorname{ch} x$ is identical with the figure of the catenary, *cf.* Chap. XIV, § 6.

Moreover: $\operatorname{sh} 0 = 0, \quad \operatorname{ch} 0 = 1, \quad \operatorname{th} 0 = 0.$

Also:

$$\operatorname{ch}^2 x - \operatorname{sh}^2 x = 1,$$

$$1 - \operatorname{th}^2 x = \operatorname{sech}^2 x, \qquad \coth^2 x - 1 = \operatorname{csch}^2 x.$$

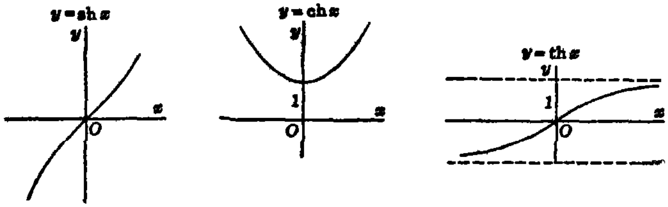


FIG. 122

The Addition Theorems are as follows:

$$\operatorname{sh}(x+y) = \operatorname{sh} x \operatorname{ch} y + \operatorname{ch} x \operatorname{sh} y;$$

$$\operatorname{ch}(x+y) = \operatorname{ch} x \operatorname{ch} y + \operatorname{sh} x \operatorname{sh} y;$$

$$\operatorname{th}(x+y) = \frac{\operatorname{th} x + \operatorname{th} y}{1 + \operatorname{th} x \operatorname{th} y}.$$

From these relations follow at once:

$$\operatorname{sh} 2x = 2 \operatorname{sh} x \operatorname{ch} x,$$

$$\operatorname{ch} 2x = \operatorname{ch}^2 x + \operatorname{sh}^2 x = 2 \operatorname{ch}^2 x - 1 = 1 + 2 \operatorname{sh}^2 x.$$

Derivatives of the Hyperbolic Functions. The derivatives have the values:

$$\frac{d \operatorname{sh} x}{dx} = \operatorname{ch} x, \qquad \frac{d \operatorname{ch} x}{dx} = \operatorname{sh} x,$$

$$\frac{d \operatorname{th} x}{dx} = \operatorname{sech}^2 x, \qquad \frac{d \coth x}{dx} = -\operatorname{csch}^2 x.$$

The Inverse Functions. The inverse of the hyperbolic sine is called the anti-hyperbolic sine:

$$y = \operatorname{sh}^{-1} x \quad \text{if} \quad x = \operatorname{sh} y.$$

Hence $\qquad x = \tfrac{1}{2}(e^y - e^{-y}).$

Solving for e^y, we get:

$$e^y = x \pm \sqrt{1 + x^2}.$$

Since $e^y > 0$ for all values of y, the upper sign alone is possible, and

$$y = \operatorname{sh}^{-1} x = \log(x + \sqrt{1 + x^2}).$$

The anti-hyperbolic cosine, however, is multiple-valued, as appears from a glance at its graph, obtained as usual in the case of an inverse function by rotating the graph of the direct function about the bisector of the angle made by the positive coordinate axes:

$$\mathrm{ch}^{-1} x = \log(x \pm \sqrt{x^2 - 1}), \qquad x \geqq 1.$$

The upper sign corresponds to positive values of $\mathrm{ch}^{-1} x$.

Also:

$$\mathrm{th}^{-1} x = \tfrac{1}{2} \log \frac{1+x}{1-x}, \qquad -1 < x < 1.$$

The derivatives have the values:

$$\frac{d\,\mathrm{sh}^{-1} x}{dx} = \frac{1}{\sqrt{1+x^2}},$$

$$\frac{d\,\mathrm{ch}^{-1} x}{dx} = \pm \frac{1}{\sqrt{x^2-1}},$$

$$\frac{d\,\mathrm{th}^{-1} x}{dx} = \frac{1}{1-x^2}.$$

We thus obtain a close analogy between certain formulas of integration:

$$\int \frac{dx}{\sqrt{a^2 - x^2}} = \sin^{-1} \frac{x}{a}, \qquad \int \frac{dx}{\sqrt{a^2 + x^2}} = \mathrm{sh}^{-1} \frac{x}{a};$$

$$\int \frac{dx}{a^2 + x^2} = \frac{1}{a} \tan^{-1} \frac{x}{a}, \qquad \int \frac{dx}{a^2 - x^2} = \frac{1}{a} \mathrm{th}^{-1} \frac{x}{a}.$$

A collection of formulas relating to the hyperbolic functions will be found in Peirce's *Tables*, pp. 81–83, and tables for $\mathrm{sh}\, x$ and $\mathrm{ch}\, x$ are given there on pp. 119–123.

Relation to the Equilateral Hyperbola. The formula:

$$\int_0^x \sqrt{1 - x^2}\, dx = \tfrac{1}{2} x \sqrt{1 - x^2} + \tfrac{1}{2} \sin^{-1} x.$$

expresses the area $OQPA$ under a circle in terms of the function $\sin^{-1} x$ and enables us, on subtracting the area of the triangle OQP from each side of the equation, to interpret $\sin^{-1} x$ as twice the area of the circular sector OPA.

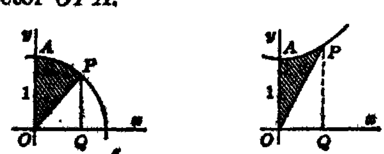


FIG. 123

There is a similar interpretation for $\operatorname{sh}^{-1} x$ with reference to the equilateral hyperbola

$$y^2 = 1 + x^2.$$

$$\int_0^x \sqrt{1+x^2}\,dx = \tfrac{1}{2}x\sqrt{1+x^2} + \tfrac{1}{2}\log(x + \sqrt{1+x^2})$$

$$= \tfrac{1}{2}x\sqrt{1+x^2} + \tfrac{1}{2}\operatorname{sh}^{-1} x.$$

Thus we see that $\operatorname{sh}^{-1} x$ is represented by twice the area of the hyperbolic sector, OPA.

To the formulas for the circle:

$$x^2 + y^2 = 1,$$

$$x = \sin u, \qquad y = \cos u,$$

correspond the following formulas for the hyperbola:

$$y^2 - x^2 = 1,$$

$$x = \operatorname{sh} u, \qquad y = \operatorname{ch} u,$$

the parameter u being represented geometrically in each case by twice the area of one of the above sectors.

The analogy of the hyperbolic functions to the trigonometric functions is but another phase of the fact that in the domain of complex quantities the trigonometric and the exponential functions and their inverse functions, the anti-trigonometric functions and the logarithms, are closely related. Compare the formulas which define $\operatorname{sh} x$ and $\operatorname{ch} x$ with those of § 7 which express $\sin x$ and $\cos x$ in terms of e^x.

The Gudermannian. Let ϕ be defined as a function of x by the relation:

$$\operatorname{sh} x = \tan\phi, \qquad \phi = \tan^{-1}\operatorname{sh} x, \qquad -\frac{\pi}{2} < \phi < \frac{\pi}{2}.$$

Then ϕ is called the *Gudermannian* of x and is denoted as follows:*

$$\phi = \operatorname{gd} x.$$

We have:

$$\operatorname{sh} x = \tan\phi, \qquad \operatorname{ch} x = \sec\phi, \qquad \operatorname{th} x = \sin\phi,$$

$$\operatorname{csch} x = \cot\phi, \qquad \operatorname{sech} x = \cos\phi, \qquad \operatorname{coth} x = \csc\phi;$$

and since $$e^x = \operatorname{ch} x + \operatorname{sh} x,$$

$$e^x = \tan\left(\frac{\pi}{4} + \frac{\phi}{2}\right), \qquad x = \log\tan\left(\frac{\pi}{4} + \frac{\phi}{2}\right).$$

* Also called the *hyperbolic amplitude* and denoted by $\operatorname{amh} x$.

EXERCISES

1. Show that

$$\sin xi = i \operatorname{sh} x, \qquad \cos xi = \operatorname{ch} x.$$

2. Show that

$$\sin(x + yi) = \sin x \operatorname{ch} y + i \cos x \operatorname{sh} y.$$

10. Limits and Continuity. A complex variable, as

$$s_n = u_1 + u_2 + \cdots + u_n,$$

where the u_k's are complex numbers, is said to *approach a limit*, if the points of the complex plane which represent it, approach a limiting point.

An infinite series of complex terms,

$$u_1 + u_2 + \cdots$$

is said to be *convergent* if the sum of its first n terms, s_n, approaches a limit.

In order that a complex variable, $Z = X + Yi$, approach a limit, it is obviously necessary and sufficient that the real part, X, by itself approach a limit, and that, at the same time, the coefficient Y of the pure imaginary part, by itself, approach a limit.

Let S be any two-dimensional region of the complex z-plane, and let a complex number, w, be assigned by any rule whatever to each point, z, of S. Then w is said to be a *function* of z, and may be written in the form

$$w = f(z).$$

The function $f(z)$ is said to be *continuous* at a point z_0 of S if $f(z)$ approaches a limit when z approaches z_0 in any manner whatever, and if the value of this limit is $f(z_0)$:

$$\lim_{z \doteq z_0} f(z) = f(z_0).$$

The foregoing definitions may be formulated as follows. The function $f(z)$ approaches a limit, A, when z approaches z_0, if, to any arbitrary positive real number, ϵ, there corresponds a positive real number δ such that

$$|f(z) - A| < \epsilon,$$

where z has any value for which

$$0 < |z - z_0| < \delta.$$

By saying that ϵ is arbitrary we mean that it may be chosen again and again, and each time, as small as we please, but real and positive.

The function $f(z)$ is *continuous* at z_0 if

$$|f(z)-f(z_0)|<\epsilon, \qquad |z-z_0|<\delta.$$

A function of a complex variable may be defined only along a curve of the complex z-plane, or even merely in a set of discrete points, like s_n.

EXERCISE

Show that, if A is any complex number such that $|A|<1$, the formula of Elementary Algebra:

$$\frac{1}{1-A}=1+A+A^2+\cdots+A^{n-1}+\frac{A^n}{1-A}$$

still holds, and thus the infinite series

$$1+A+A^2+\cdots$$

converges to the value $1/(1-A)$.

11. The Derivative. The definition of a derivative given at the beginning of the Calculus holds here unmodified. Let z_0 be any point of S, and let w_0 be the corresponding value of the function:

$$w_0=f(z_0).$$

Give to z an increment, Δz, merely such that $z_0+\Delta z$ is a point of S, and denote the value of the function by $w_0+\Delta w$:

$$w_0+\Delta w=f(z_0+\Delta z).$$

Form the difference-quotient:

$$\frac{\Delta w}{\Delta z}=\frac{f(z_0+\Delta z)-f(z_0)}{\Delta z}.$$

If this variable approaches a limit when Δz approaches 0, then $f(z)$, or w, is said to have a *derivative* at $z=z_0$, and we write:

$$\lim_{\Delta z\doteq 0}\frac{\Delta w}{\Delta z}=D_zw \qquad \text{or} \qquad \lim_{\Delta z\doteq 0}\frac{f(z_0+\Delta z)-f(z_0)}{\Delta z}=f'(z_0).$$

The five General Theorems, Theorems I–V, *Introduction to the Calculus*, pp. 22–35, hold as in the case of reals:

I. $$D_z(cw)=cD_zw,$$

II. $$D_z(w_1+w_2)=D_zw_1+D_zw_2,$$

etc.

Moreover, the differential is defined in the same manner:

$$dw = D_z w \cdot \Delta z,$$

and it is shown that

$$dw = D_z w\, dz,$$

even when both z and w depend on a third complex variable, t.

The special formulas:

$$dc = 0, \qquad d\,z^n = n z^{n-1}\, dz, \quad n, \text{ a natural number,}$$

are also established as in the real case of reals.

It follows at once, as in the case of reals, that all polynomials have derivatives. Thus

$$\frac{d}{dz}(az^2 + bz + c) = 2az + b,$$

where the coefficients a, b, c, are any complex constants.

Also, any rational function,

$$R(z) = \frac{g(z)}{G(z)},$$

has a derivative for any value of z for which it is defined (*i.e.* $G(z) \neq 0$). In the case of transcendental functions, however, the proof of the existence of a derivative is indirect, and will be taken up in the next paragraph.

12. The Cauchy-Riemann Differential Equations. A very simple function of a complex variable may fail to have a derivative. Consider, for example,

$$w = x - yi.$$

This is a function of z, for, when z is given, w is determined; and moreover the function is obviously continuous. Give to z a value z_0 and form the difference-quotient:

$$w_0 = x_0 - iy_0, \qquad w_0 + \Delta w = (x_0 + \Delta x) - i\,(y_0 + \Delta y),$$

$$\frac{\Delta w}{\Delta z} = \frac{\Delta x - i\Delta y}{\Delta x + i\Delta y}.$$

First, let Δz approach the limit 0, passing through real values. Then $\Delta y = 0$ and

$$\frac{\Delta w}{\Delta z} = 1.$$

Hence $\Delta w/\Delta z$ has the limit 1 for this mode of approach.

Next, let Δz approach 0, passing through values that are pure imaginaries. Then $\Delta x = 0$ and

$$\frac{\Delta w}{\Delta z} = -1.$$

Here, $\Delta w/\Delta z$ has the limit -1. Thus no limit exists when Δz approaches 0, for the points of the complex plane which represent the variable $\Delta w/\Delta z$ do not lie near some single fixed point when Δz lies near 0.

Necessary Conditions. It is easy to obtain a pair of necessary conditions * that the function

$$w = f(z)$$

have a derivative. Let $w = u + vi$, and let Δz approach 0, passing through real values: $\Delta y = 0$, $\Delta z = \Delta x$. Then

$$\lim_{\Delta z \doteq 0} \frac{\Delta w}{\Delta z} = \lim_{\Delta x \doteq 0} \left(\frac{\Delta_x u}{\Delta x} + i \frac{\Delta_x v}{\Delta x} \right) = \frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x},$$

and hence

$$D_z w = \frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x}.$$

Next, let Δz approach 0, passing through pure imaginary values: $\Delta x = 0$, $\Delta z = i \Delta y$. Then

$$\lim_{\Delta z \doteq 0} \frac{\Delta w}{\Delta z} = \lim_{\Delta y \doteq 0} \left(\frac{1}{i} \frac{\Delta_y u}{\Delta y} + \frac{\Delta_y v}{\Delta y} \right) = \frac{1}{i} \frac{\partial u}{\partial y} + \frac{\partial v}{\partial y} = \frac{\partial v}{\partial y} - i \frac{\partial u}{\partial y}.$$

Hence

$$D_z w = \frac{\partial v}{\partial y} - i \frac{\partial u}{\partial y}.$$

On equating these two values of $D_z w$ and observing that the real parts of the two expressions must be equal by themselves, we see that

$$(A) \qquad \frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \qquad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}.$$

The result may also be written in the form of a single equation in complex quantities:

$$(A') \qquad \frac{\partial w}{\partial x} = \frac{1}{i} \frac{\partial w}{\partial y}.$$

These are the necessary conditions we set out to obtain.

* The student will do well to turn back to the *Introduction to the Calculus*, p. 71, and make sure that he has clearly in mind what is meant by a *necessary*, and what is meant by a *sufficient* condition.

Sufficient Conditions. If we assume that the real functions $u = u(x, y)$ and $v = v(x, y)$ are continuous, together with their first partial derivatives, then the equations (A) form, conversely, a *sufficient condition*, that $w = u + vi$ have a derivative. For,

$$\Delta u = \frac{\partial u}{\partial x}\Delta x + \frac{\partial u}{\partial y}\Delta y + \zeta_1 \Delta x + \zeta_2 \Delta y,$$

$$\Delta v = \frac{\partial v}{\partial x}\Delta x + \frac{\partial v}{\partial y}\Delta y + \zeta_1' \Delta x + \zeta_2' \Delta y,$$

where the ζ's are all infinitesimals, Δx and Δy being the principal infinitesimals; Chap. V, § 3. Form the quotient,

$$\frac{\Delta w}{\Delta z} = \frac{\Delta u + i\,\Delta v}{\Delta x + i\,\Delta y}.$$

Since equations (A) are true, it is clear that

$$\left(\frac{\partial u}{\partial x}\Delta x + \frac{\partial u}{\partial y}\Delta y\right) + i\left(\frac{\partial v}{\partial x}\Delta x + \frac{\partial v}{\partial y}\Delta y\right) = \left(\frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x}\right)(\Delta x + i\,\Delta y).$$

Hence

$$\frac{\Delta w}{\Delta z} = \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x} + (\zeta_1 + i\,\zeta_1')\frac{\Delta x}{\Delta z} + (\zeta_2 + i\,\zeta_2')\frac{\Delta y}{\Delta z}.$$

Now $|\Delta x| \leqq |\Delta z|$ and $|\Delta y| \leqq |\Delta z|$, and the parentheses, $\zeta_1 + i\zeta_1'$ and $\zeta_2 + i\zeta_2'$, are infinitesimal. Hence $\Delta w/\Delta z$ approaches a limit when Δz approaches 0, and thus the function w is seen to possess a derivative. Moreover, the value of the derivative is given by the formula:

$$(B) \qquad D_z w = \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x}.$$

The equations (A) are known as the *Cauchy-Riemann Differential Equations.*

A function w of the complex variable z which possesses a continuous derivative is called an *analytic function.* Unlike the situation in the case of reals, it is only the analytic functions of a complex variable which have important properties.

Example. The function e^z is analytic. Here,

$$u = e^x \cos y, \qquad v = e^x \sin y.$$

$$\frac{\partial u}{\partial x} = e^x \cos y, \qquad \frac{\partial v}{\partial x} = e^x \sin y,$$

$$\frac{\partial u}{\partial y} = -e^x \sin y, \qquad \frac{\partial v}{\partial y} = e^x \cos y,$$

and the Cauchy-Riemann Differential Equations (A) are satisfied. Thus e^z has a derivative, and it is given by the formula:

$$\frac{\partial u}{\partial x}+i\frac{\partial v}{\partial x}=e^x\cos y+i\,e^x\sin y=e^z.$$

Hence

(1) $$\frac{d}{dz}e^z=e^z.$$

EXERCISES

1. Show that the function $\sin z$ has a derivative given by the formula:

$$\frac{d\sin z}{dz}=\cos z.$$

2. Show similarly that

$$\frac{d\cos z}{dz}=-\sin z;\qquad \frac{d\tan z}{dz}=\sec^2 z;\qquad \frac{d\cot z}{dz}=-\csc^2 z.$$

3. Prove that the Cauchy-Riemann Differential Equations, transformed to polar coordinates, are as follows:

$$\frac{\partial u}{\partial r}=\frac{1}{r}\frac{\partial v}{\partial \theta},\qquad \frac{1}{r}\frac{\partial u}{\partial \theta}=-\frac{\partial v}{\partial r}.$$

4. Show that the function

$$w=\log z$$

has a derivative, and that

$$\frac{d\log z}{dz}=\frac{1}{z}.$$

5. Prove that

$$\frac{d\sin^{-1}z}{dz}=\frac{1}{\sqrt{1-z^2}},\qquad \frac{d\tan^{-1}z}{dz}=\frac{1}{1+z^2}.$$

13. Laplace's Equation, $\Delta u=0$. If $w=u+vi$ is an analytic function of the complex variable $z=x+yi$, then the real part, u, of the function satisfies Laplace's Equation:

(1) $$\Delta u\equiv\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=0.$$

For, u and v satisfy the Cauchy-Riemann Differential Equations

(*A*), § 12. On differentiating* the first of these with respect to x, the second with respect to y, and adding, equation (1) results.

Conversely, let u be a solution of Laplace's Equation, (1), throughout a region, S, which, for simplicity, we assume simply connected. Form the integral

$$v = \int_{(a,b)}^{(x,y)} -\frac{\partial u}{\partial y}\,dx + \frac{\partial u}{\partial x}\,dy.$$

This integral is of the form

$$\int_{(a,b)}^{(x,y)} P\,dx + Q\,dy.$$

Moreover, the condition (*cf.* Chap. XI, § 7)

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$$

here reduces to

$$-\frac{\partial^2 u}{\partial y^2} = \frac{\partial^2 u}{\partial x^2},$$

and thus is fulfilled because of (1). Hence the integral is independent of the path of integration and thus represents a function v, single-valued in S.

Furthermore, the function v thus defined has partial derivatives given by the formulas:

$$\frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}, \qquad \frac{\partial v}{\partial y} = \frac{\partial u}{\partial x}.$$

But these are precisely the Cauchy-Riemann Differential Equations (*A*), § 12. Hence u is the real part of an analytic function $w = u + vi$ of the complex variable $z = x + yi$. The functions u and v are called *conjugate*; *cf.* Chap. XI, § 17.

Thus the theory of Laplace's Equation in two dimensions is coextensive with the theory of analytic functions of a complex variable.

14. Cauchy's Integral Theorem. By the *integral* of a continuous function of a complex variable along a curve C is meant the limit of the sum:

* It can be shown that any two functions, u and v, which are continuous, together with their first partial derivatives, and satisfy (*A*), § 12, possess continuous second partial derivatives.

(1) $$\lim_{n=\infty}\sum_{k=0}^{n-1} f(z_k)\,\Delta z_k = \int_C f(z)\,dz,$$

where $\Delta z_k = z_{k+1} - z_k$. Since

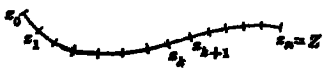


FIG. 124

$$f(z_k)\,\Delta z_k = (u_k + i v_k)(\Delta x_k + i\Delta y_k)$$
$$= u_k\,\Delta x_k - v_k\,\Delta y_k + i\,(v_k\,\Delta x_k + u_k\,\Delta y_k),$$

it follows that

(2) $$\int_C f(z)\,dz = \int_C u\,dx - v\,dy + i\int_C v\,dx + u\,dy.$$

Let M be the largest value which $|f(z)|$ takes on along C. Since by § 3, (2),

$$\left|\sum_{k=0}^{n-1} f(z_k)\,\Delta z_k\right| \leqq \sum_{k=0}^{n-1} |f(z_k)|\cdot|\Delta z_k|,$$

and since $|\Delta z_k| = l_k$, the length of the chord joining z_k with z_{k+1}, it follows that the last sum is not greater than

$$M(l_0 + l_1 + \cdots + l_{n-1}).$$

The parenthesis, being the length of a broken line inscribed in C, approaches as its limit the length, l, of C. Hence we infer that

$$\left|\int_C f(z)\,dz\right| \leqq Ml.$$

CAUCHY'S INTEGRAL THEOREM. *Let $f(z)$ be analytic throughout the interior of a region S and continuous on the boundary, C. Then*

$$\int_C f(z)\,dz = 0,$$

where the integral is extended in the positive sense over the entire boundary, C.

For, each of the line integrals on the right of (2) is of the form

$$\int_C P\,dx + Q\,dy,$$

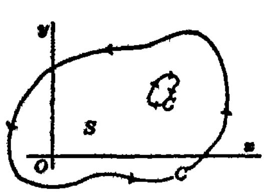


FIG. 125

and since $f(z)$ is analytic, u and v satisfy the Cauchy-Riemann Differential Equations (*A*), § 12. Hence the further condition,

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x},$$

is fulfilled for each integral, and so each vanishes; Chap. XI, § 7.

15. Cauchy's Integral Formula. From the theorem of § 14 Cauchy deduced the following formula:

$$f(z) = \frac{1}{2\pi i}\int_C \frac{f(t)\,dt}{t-z}, \tag{1}$$

where z is any interior point of S, and the integral is extended over the entire boundary C in the positive sense.

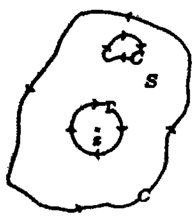


FIG. 126

Draw a small circle, Γ, about z and remove this circle from S. In the region S' thus obtained,

$$\frac{f(t)}{t-z},$$

regarded as a function of t, z being constant, satisfies the conditions of the theorem of § 14. Hence

$$\int_C \frac{f(t)\,dt}{t-z} + \int_\Gamma \frac{f(t)\,dt}{t-z} = 0. \tag{2}$$

The second integral is extended in the clockwise sense, and can be evaluated as follows. Let

$$t - z = \rho e^{\theta i},$$

where t is a point of Γ, and ρ is the radius. Then

$$dt = i\rho e^{\theta i}\,d\theta$$

and

$$\int_\Gamma \frac{f(t)\,dt}{t-z} = \int_{2\pi}^{0} \frac{f(t)\,i\rho e^{\theta i}\,d\theta}{\rho e^{\theta i}} = -\,i\int_0^{2\pi} f(z+\rho e^{\theta i})\,d\theta.$$

Now let ρ approach the limit 0. The last integral approaches

$$\int_0^{2\pi} f(z)\,d\theta = 2\pi f(z);$$

for the integrand is continuous in the two independent variables, ρ and θ. The same is, therefore, true of its real and its pure imaginary part, and so the theorem of Chap. XIX, § 1 can be applied to each of these.

The first term in (2) does not depend on ρ. Hence we have:

$$\int_C \frac{f(t)\,dt}{t-z} - 2\pi i f(z) = 0,$$

and thus (1) is established.

Differentiation under the Sign of Integration. An integral of the above form:

$$\phi(z)=\int_C \frac{f(t)\,dt}{t-z},$$

where $f(t)$ is continuous along C, can be differentiated according to Leibniz's Rule. For,

$$\phi(z+\Delta z)-\phi(z)=\int_C f(t)\left(\frac{1}{t-z-\Delta z}-\frac{1}{t-z}\right)dt.$$

Hence

$$\frac{\phi(z+\Delta z)-\phi(z)}{\Delta z}=\int_C \frac{f(t)\,dt}{(t-z-\Delta z)(t-z)},$$

and

$$\frac{\phi(z+\Delta z)-\phi(z)}{\Delta z}-\int_C \frac{f(t)\,dt}{(t-z)^2}=\Delta z\int_C \frac{f(t)\,dt}{(t-z-\Delta z)(t-z)^2}.$$

This last integral remains finite as Δz approaches 0. The right-hand side of the equation, therefore, approaches 0, and the theorem is proved:

(3) $$\phi'(z)=\int_C \frac{f(t)\,dt}{(t-z)^2}.$$

From (3) we infer the following formula for $f'(z)$:

(4) $$f'(z)=\frac{1}{2\pi i}\int_C \frac{f(t)\,dt}{(t-z)^2}.$$

The process can be repeated indefinitely, and thus we have:

(5) $$f^{(n)}(z)=\frac{n!}{2\pi i}\int_C \frac{f(t)\,dt}{(t-z)^{n+1}}.$$

EXERCISE

If S is a circle with centre $z=a$ and radius r, and if M is an upper limit for $|f(t)|$ on the circumference, C:

$$|f(t)|_C \leqq M,$$

show that

$$\left|\int_C \frac{f(t)\,dt}{t-a}\right| \leqq \int_0^{2\pi} \frac{Mr\,d\theta}{r}=2\pi M.$$

Hence $$|f(a)| \leqq M.$$

Show further that
$$|f^{(n)}(a)| \leqq Mn!\,r^{-n}.$$

16. Taylor's Theorem. Let $f(z)$ be analytic within a region S and continuous on the boundary, C. Let a be any interior point of S and let r be the radius of the largest circle about a which contains no point of C in its interior. Then $f(z)$ can be developed by Taylor's Theorem throughout the interior of this circle:

$$(1) \qquad f(z) = f(a) + f'(a)(z-a) + \frac{f''(a)}{2!}(z-a)^2 + \cdots, \quad |z-a| < r.$$

The proof is brief. Let t be any point of C, and let z be an interior point of the circle. Then

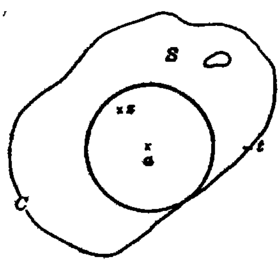


FIG. 127

$$|z-a| < r, \qquad |t-a| \geqq r.$$

Write

$$\frac{1}{t-z} = \frac{1}{t-a-(z-a)} = \frac{1}{t-a}\cdot\frac{1}{1-\dfrac{z-a}{t-a}}.$$

This last fraction is of the form $1/(1-A)$, where $|A| < 1$. The formula of Elementary Algebra for the sum in a geometric progression holds here, § 10, Exercise, and thus we have:

$$\frac{1}{t-z} = \frac{1}{t-a} + \frac{z-a}{(t-a)^2} + \cdots + \frac{(z-a)^{n-1}}{(t-a)^n} + \frac{(z-a)^n}{(t-z)(t-a)^n}.$$

It remains merely to multiply this equation through by $f(t)/2\pi i$, to integrate over C, and to interpret the terms:

$$\frac{1}{2\pi i}\int_C \frac{f(t)\,dt}{t-z} = \sum_{k=0}^{n-1}\frac{(z-a)^k}{2\pi i}\int_C \frac{f(t)\,dt}{(t-a)^{k+1}} + \frac{1}{2\pi i}\int_C \frac{(z-a)^n f(t)\,dt}{(t-z)(t-a)^n}.$$

The term on the left is equal to $f(z)$ by § 15, (1), and the n terms of the sum on the right are precisely the first n terms of the Taylor's Series, by § 15, (1), (4), and (5). Since

$$\left|\int_C \frac{(z-a)^n f(t)\,dt}{(t-z)(t-a)^n}\right| \leqq \int_C \left|\frac{z-a}{t-a}\right|^n \frac{|f(t)|}{|t-z|}\,ds$$

and $\left|\dfrac{z-a}{t-a}\right| \leqq \dfrac{|z-a|}{r} = h$, a positive constant < 1; since furthermore $|t-z| \geqq r - |z| = g$, a positive constant; and since finally $|f(t)| \leqq M$, a positive constant, the last term on the right is in absolute value not greater than

$$\frac{Ml}{2\pi g}h^n,$$

where l denotes the total length of C.

As n increases without limit, this quantity approaches 0. Hence the series (1) converges and represents the function $f(z)$.

If, in particular, $f(z)$ is analytic over the entire complex plane, the series (1) converges for all values of z and represents the function. Thus the developments

$$e^z = 1 + z + \frac{z^2}{2!} + \cdots,$$

$$\sin z = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \cdots,$$

$$\cos z = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \cdots$$

are seen to hold for all values of z, real or complex.

Taylor's Theorem throws light on the extent of the interval of convergence of an expansion of a real function into a power series. Take, for example, the series which represents the function $1/(1+x^2)$:

(2) $$\frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + \cdots.$$

This function is continuous, together with all its derivatives, for all real values of x, and can be expanded about any real point, $x = a$. Why should the above series converge in just the interval

$$-1 < x < 1?$$

The question is answered by considering the function

(3) $$\frac{1}{1+z^2}$$

in the complex plane. This function has singular points at $z = i$, $-i$. Hence the largest circle which can be drawn about the origin ($z = 0$) and which contains in its interior only points in which the function is analytic, is the unit circle, and this circle cuts off from the x-axis the interval in question.

It should be added for completeness that a power series in z which converges for some values of the argument distinct from 0 and diverges for others, always converges throughout the interior of a certain circle whose centre is at $z=0$, and diverges outside this circle. It represents, or defines, an analytic function within its circle of convergence.

Finally, Taylor's expansion is unique; *i.e.* no second expansion into a power series is possible, whose coefficients are different from those of the Taylor's series.

The proof of Taylor's Theorem was given by Cauchy in 1831.

EXERCISES

1. Show that the function (2) can be expanded about the point $x=1$:

$$\frac{1}{1+x^2}=c_0+c_1(x-1)+c_2(x-1)^2+\cdots.$$

Determine the first three coefficients, and prove that the interval of convergence of the series is

$$1-\sqrt{2}<x<1+\sqrt{2}.$$

2. Show that the function

$$\frac{1}{1+x+x^2}$$

can be developed into a power series in x, convergent throughout the interval $-1<x<1$, and determine the first three coefficients.

3. Show that $\tan x$ can be developed into a power series in x, convergent in the interval

$$-\frac{\pi}{2}<x<\frac{\pi}{2}.$$

4. Show that the function

$$\frac{1+x^2}{5-x+x^2}$$

can be developed about the point $x=2$, and determine the interval of convergence of the power series.

5. Show that any rational function,

$$R(z)=\frac{g(z)}{G(z)},$$

can be developed by Taylor's Theorem about any point, a, at which $G(a)\neq 0$, and that the circle of convergence will reach out to the nearest root of $G(z)$.

17. Multiple-Valued Functions. When $f(z)$ is multiple-valued, it is possible in the cases which arise in practice to separate out a *branch* which is single-valued and analytic throughout a certain region. Consider, for example,

$$\frac{1}{\sqrt{1-2\mu z+z^2}}, \qquad |\mu| \leqq 1. \tag{1}$$

Here, the radicand vanishes when

$$1-2\mu z+z^2=0, \qquad z=\mu \pm i\sqrt{1-\mu^2}.$$

If μ is real, these two points lie on the unit circle. Within this circle, the radicand is never 0, and the two values of the function can be grouped so as to give *two* functions, each analytic within this circle. Let $f(z)$ be that one of these functions for which $f(0)=1$. Then $f(z)$ can be developed by Taylor's Theorem about the point $z=a=0$, and we have:

$$f(z)=\frac{1}{\sqrt{1-2\mu z+z^2}}=P_0(\mu)+P_1(\mu)z+P_2(\mu)z^2+\cdots. \tag{2}$$

The coefficients, $P_m(\mu)$, can be shown to be the Zonal Harmonics.

EXERCISES

1. Compute the first three coefficients of the expansion (2) by differentiation, and compare the results with the formulas of Chap. XVI, § 5.

2. Show that the function

$$\log \cos x$$

can be developed into a Maclaurin's series, and determine the interval of convergence.

18. Conformal Mapping. We have pointed out in Chap. VI, § 8, the nature of a conformal map of one surface on another. If a region S of the (x, y)-plane is mapped in a one-to-one manner on a region Σ of the (u, v)-plane by the functions

$$u=f(x, y), \qquad v=\phi(x, y),$$

assumed continuous together with their first partial derivatives, and if the angle under which any two curves in S intersect, is preserved by their images in Σ, then a small triangle in S will go over into a small curvilinear triangle having respectively the same angles, and thus any small figure in S will go over into a figure in Σ which will appear similar, though drawn to a different scale and turned through an angle.

Let $$w = f(z)$$

be a function which is analytic throughout a region S of the complex plane, and let

$$\left.\frac{dw}{dz}\right|_{z=z_0} = f'(z_0) \neq 0.$$

It can be shown that the region about z_0 will be mapped in a one-to-one manner on the region about w_0 in the w-plane. Moreover, *this map is conformal.*

To prove this last statement, consider an arbitrary curve, C, emanating from z_0 and making an angle ϕ with the positive axis of reals. Its image, C', will be a curve in the w-plane emanating from w_0 and making with the positive axis of reals an angle which we will call ψ. We will show that

(1) $$\psi = \phi + \gamma,$$

where γ is the same for all curves C.

Let $z' = z_0 + \Delta z$ be a second point on C near z_0, and let $w' = w_0 + \Delta w$ be its image on C'. Since

$$\lim_{\Delta z \doteq 0} \frac{\Delta w}{\Delta z} = (D_z w)_{z=z_0} = \mathsf{A} e^{\gamma i} \neq 0,$$

we have:

$$\frac{\Delta w}{\Delta z} = \mathsf{A} e^{\gamma i} + \zeta, \qquad \Delta w = (\mathsf{A} e^{\gamma i} + \zeta)\,\Delta z,$$

where ζ is a (complex) infinitesimal.

We know from § 2 that the angle of the product of two complex quantities is the sum of the angles of the factors. Hence

$$\text{arc}\,\Delta w = \text{arc}\,(\mathsf{A} e^{\gamma i} + \zeta) + \text{arc}\,\Delta z.$$

As Δz approaches 0, the terms on the right approach respectively γ and ϕ, and the term on the left approaches ψ. Thus the truth of (1) is established.

If, now, C_1 and C_2 are two curves in the z-plane emanating from $z = z_0$ and making angles ϕ_1 and ϕ_2 with the axis of reals; and if their images, C_1' and C_2', make angles ψ_1 and ψ_2 with the axis of reals in the w-plane, then

$$\psi_1 = \phi_1 + \gamma, \qquad \psi_2 = \phi_2 + \gamma$$

and hence

$$\psi_2 - \psi_1 = \phi_2 - \phi_1;$$

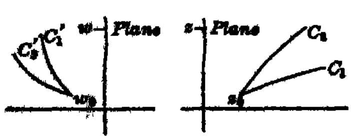


FIG. 128

i.e. the angle at w_0 from C_1' to C_2' is the same as the angle at z_0 from C_1 to C_2.

The student will find a number of carefully drawn plates representing such maps in Clerk Maxwell's *Electricity and Magnetism*, vol. I, end; and also in Holzmüller's *Isogonale Verwandtschaften*.

19. Flow of Heat or Electricity. Irrotational Fluid Motion. The two-dimensional flow of these substances has already been mentioned in Chap. XI, §§ 16, 19. It is Laplace's Equation and the Cauchy-Riemann Differential Equations which form the common basis of that great branch of Mathematical Physics and of those further developments in analysis on which higher mathematics rests — the Theory of Functions of a Complex Variable.

NOTE ON CHAP. VII, § 5, *Example*, p. 181.

It is possible to generalize this example and at the same time to simplify the treatment. Let $F(x, y, z)$ be any quadratic form whatsoever. Then F has a maximum on the sphere $\Phi = a^2$. Hence Equations (11) must hold at this point for a suitable value of λ. Let the axes of coordinates be so rotated that the point in question is $(0, 0, a)$. Then it follows from (11) that, for the transformed equations,

$$D = 0, \qquad E = 0.$$

Thus

$$F(x, y, z) = Ax^2 + 2Fxy + By^2 + Cz^2$$

If the term in xy is present, a suitable rotation of the axes about the axis of z will remove it (*Analytic Geometry*, Chap. XII, § 2) and thus $F(x, y, z)$ is reduced to the form (13). But now A, B, C can be any real numbers whatsoever.

Thus the possibility of reducing a quadratic form to a sum of squares by means of rotations is shown. The determination of the actual transformations which will yield the result is a question of less importance, though easily answered by Equations (11). It is not our purpose to treat it here.

INDEX